# A RELATIONAL VIEW OF CONSUMER SOCIAL LEARNING IN VIRTUAL COMMUNITIES

by

Bo Liang

A dissertation submitted to the faculty of
The University of Utah
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in

Business Administration

David Eccles School of Business

The University of Utah

December 2014

**The University of Utah Graduate School**

# STATEMENT OF DISSERTATION APPROVAL

The dissertation of **Bo Liang**

has been approved by the following supervisory committee members:

| | | |
|---|---|---|
| **Debra L. Scammon** | , Chair | **July 25, 2014** Date Approved |
| **Abbie Griffin** | , Member | **July 25, 2014** Date Approved |
| **Stephen J. Carson** | , Member | **July 25, 2014** Date Approved |
| **Shyam Gopinath** | , Member | **July 25, 2014** Date Approved |
| **Vandana Ramachandran** | , Member | **July 25, 2014** Date Approved |

and by **William Hesterly** , Associate Dean

of the **David Eccles School of Business**

and by David B. Kieda, Dean of The Graduate School.

# ABSTRACT

Social learning is an important part of consumers' lives. A major limitation of previous research in consumer social learning is the lack of attention to the effects of relationships built among participants on the learning process in online communities. Within a thread in an online forum, when a post is made referring to a previous post, a connection is built between the posts, indicating the opportunity for the continuation of the learning process. We construct a conceptual framework of reference relationships in ongoing threads. We propose that four key constructs related to the reference relationships between posts are associated with the continuation of the social learning process within threads, including engagement, sociality, advanced levels of learning, and existence and strength of ties between posters. We investigated 19 threads with 580 posts in a diabetes online community using both qualitative and quantitative analysis. Results provide substantial evidence that reference relationships in an ongoing thread are significantly positively associated with posters' engagement in the community and in the ongoing thread, advanced learning steps and sociality demonstrated in the content of posts in the ongoing thread, and ties in other threads. Our study suggests important implications for marketers in facilitating consumers' social learning process in online communities. Specifically, we suggest that by investigating posts made in reference relationships, marketers may identify the topics in which posters have interest and the influentials who disseminate and generate knowledge in online communities.

This dissertation is dedicated to my parents, Yumin Liang and Gaifeng Guo,

for their endless love and support.

TABLE OF CONTENTS

# ACKNOWLEDGEMENTS

I would like to express my special appreciation and thanks to my advisor Dr. Debra Scammon, who has been a tremendous mentor for me. I would like to thank her for encouraging my research and for allowing me to grow as a researcher. Her advice on both research as well as on my career have been priceless. I would also like to thank my committee members, Dr. Abbie Griffin, Dr. Steve Carson, Dr. Shyam Gopinath, and Dr. Vandana Ramachandran, for serving as my committee members. I also want to thank them for letting my defense be an enjoyable moment and for their brilliant comments and suggestions. Lastly I would like to give special thanks to my family. Words cannot express how grateful I am to my mother, my husband, and my son for all of the sacrifices that they have made on my behalf and all of the support that they have given me to achieve my goals.

# INTRODUCTION

Social learning theory suggests that people learn in their social contexts (Bandur, 1971). Social learning is an important part of consumers' lives. Through social learning, consumers gain vicarious experience with using products and learn from others. Previous research has found that social learning has significant effects on consumer knowledge generation and decision making. In recent years, researchers in consumer social learning have paid significant attention to social learning in online communities.

As online communication becomes a large part of consumers' everyday life, online social learning is attracting a great deal of attention among both academics and practitioners. Consumers are actively engaged in sharing their experiences and views about products or services with others via social media. Learning from others in online communities can increase consumer product knowledge and assist consumers in exploring potential benefits from products and consumptions behaviors.

Virtual communities in which an individual can communicate with "friends" or informed others (Ellison, 2007) are becoming a common means for health communication. Nearly half of internet users who have searched for online health information reported that they have participated in online health–related communication (reading or posting messages; Fox & Jones, 2009). Research on online health communication has greatly expanded in recent years. The major findings in this field have primarily focused on (1) the benefits that consumers may acquire in online have

primarily focused on (1) the benefits that consumers may acquire in online communities such as making empowered decisions (Donelle & Hoffman-Goetz 2008; Eichhorn, 2008; Jayanti & Singh, 2009) and (2) the types of content of communication (e.g., disease-specific guidance and feedback; Greene, Choudhry, Kilabuk, & Shrank, 2010) through online health communities. Little research exists on the dynamics of relations demonstrated within threads and their importance in the consumer learning process. Based on extant findings and gaps in the research on online health communities, our study is set within online health communities.

The essential feature of social learning is that people learn within social structures. Research on organizational learning has shown the significant effects of social relations or social networks on organizational learning (Argote, 2013; Borgatti & Cross, 2003; Cross & Sproull, 2004; Hansen, Mors, & Lovas, 2005; Kraatz, 1998; Levin & Cross, 2004). Previous research on online consumer social learning has primarily addressed two questions: know-what and know-how (e.g., Greene et al., 2010; Jayanti & Singh, 2009). By analyzing the content of posts, Jayanti and Singh (2009) suggest that consumers experienced three steps with regard to social learning: identifying problems, acquiring knowledge (know-what), and taking actions (know-how). A major limitation of this line of research is the lack of attention to the relationships built among participants in online communities, a question of know-who.

Following Walther (1992,1996), we view social learning in online communities as a process in which participants not only learn from each other but also build relationships with one another through interactional activities such as making posts. Within online communities, participants create threads and within threads, participants can make a post.

When a post is made quoting/referring to a previous post, a connection is built between the poster of the current post and that of the referenced post. Previous research in communications and education in online contexts has used message reference analysis to explore the role of the reference relationships of posts in facilitating online learning (Ahern, Peck, & Laycock, 1992; Bullen, 2007; Ingram & Hathorn, 2004; Pena-Shaff & Nicholls, 2004; Sze, 2008; Wise, Hamman, & Thorson, 2006).

Through observation of several health-related forums, we found that reference relationships between posts vary across threads. We show two examples of threads: thread 1 (Figure 1) and thread 2 (Figure 2). In the figures, each rectangular box represents one post with the poster identified by a capital letter; each line with an arrow represents a reference relationship between two posts. In both threads, the first post was made by the thread initiator, poster A. Thread 1 consists of seven posts. In that thread the first and only post by the thread initiator, "A," was referenced by five posts by five unique posters: "B," "C," "D," "E," and "F," and the post by poster "B" was referenced in the post by poster "G." In this thread, one-time and unidirectional relationships were built between the initiator "A" and five other posters and between posters "B" and "G." Thread 2 consists of a total of 26 posts and demonstrates a more complex pattern of relationships between posts than does thread 1. Thread 2 includes three clusters of posts, each suggesting a continuous learning process with a number of posts connected with one another: cluster 1 with six connected posts, cluster 2 with 10 connected posts, and cluster 3 with five connected posts. In each cluster, posts are connected by being referenced in other posts or referencing to other posts. These clusters reveal multiple-time and two-directional referencing of posts between posters such as posters "A" and "F" in cluster 1

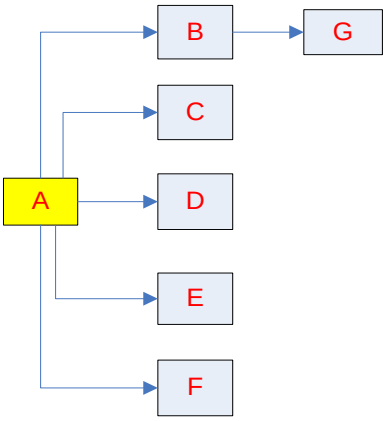


Figure 1. Thread 1

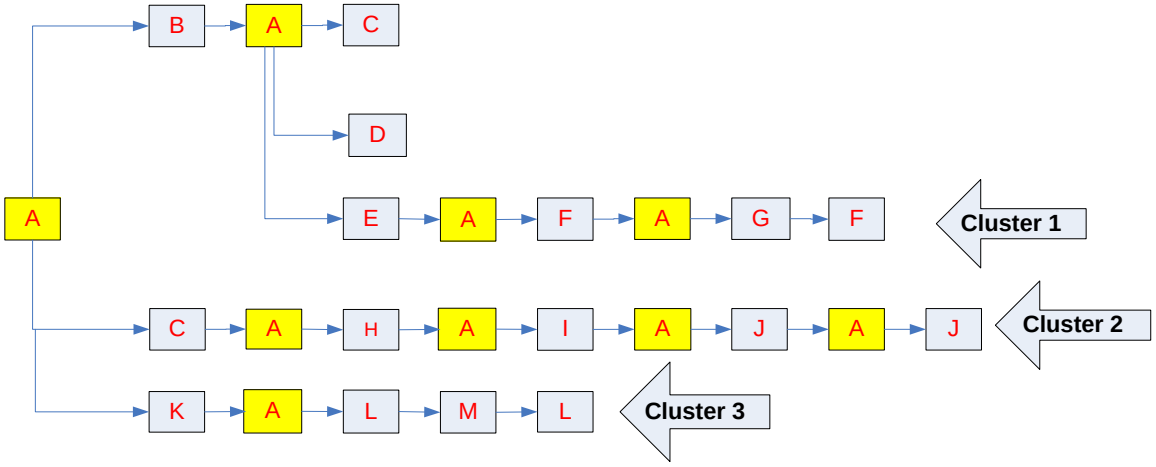


Figure 2. Thread 2

and posters “A” and “J” in cluster 2. These represent stronger connections/relationships.

Social learning processes are “cyclically connected and socially construed” (Jayanti & Singh, 2009, p. 1061). People identify problems, generate knowledge, and then take actions. While taking actions, they may identify new problems. Therefore, a new round of the learning process may begin. During this cyclical learning process, individuals interact and develop relationships with one another. With regard to online forums, it is important to determine the role that the reference relationship has in social learning. Some posts are referenced in other posts, while others are not. Whether an individual post is referenced may facilitate or inhibit the building of relationships among users in online forums. Referencing indicates a direct dialogue between the poster and the referenced poster. By referencing, the poster builds a connection with the referenced poster, facilitating the development of a relationship between the two posters. Otherwise, the poster loses an opportunity of connecting with the other poster. Additionally, it is important to determine what drives some consumers to make multiple and reciprocal posts to others, thus building strong ties with them, whereas some consumers make only a single nonreciprocal post to others, establishing weak ties with them. Without knowing the mechanism underlying the relationship between posters, it is hard to suggest strategies to facilitate consumer social learning in online forums. This study aims to examine the factors associated with social learning within threads demonstrated by reference relationships between posters.

To provide insights into relationship development among consumers and thus into opportunities for social learning in an online environment, we construct our theory based on theories of pragmatic learning (Jayanti & Singh, 2009), online communities of

consumption (Kozinets, 1997, 1998, 1999; Muniz & O'Guinn, 2005), and tie strength. Granovetter first proposed the theory of tie strength in the context of offline social networks in 1973. We extend the notion of tie strength by contextualizing our study in an online community focusing on diabetes treatments. We explore the factors that are associated with the development of reference relationships in ongoing threads.

Our study makes both theoretical and managerial contributions. We develop a conceptual framework that incorporates relationship measures into the traditional understanding of consumer social learning. Previous studies have used the number of friends listed on an individual user's friend list to measure the social influence of this user on other members in online communities. However, a user only interacts with a few friends. Our study investigates the actual interaction between users demonstrated by reference relationships. We then bridge the gap between theory (tie strength theory) and practice (reference relationships of posts) and suggest strategies for social media marketing. Quoting is a common function in online communities allowing users to "quote" any previous post. Our findings show that posts involved in reference relationships (being quoted by others or quoting others) include more advanced learning content. By mapping the reference relationships between posters, marketers may be able to identify the topics in which posters are interested and the influentials who disseminate and generate knowledge.

We begin by reviewing the relevant literature. Thereafter, we develop concepts and construct hypotheses. We then explore an online diabetes forum to illustrate the dynamics of relational communication in the social learning process. We conclude with a discussion of our results for consumer learning and implications for consumer research.

# THEORETICAL BACKGROUND

## The Knowledge Element of Social Learning

Social learning is an approach to knowledge sharing in social contexts. Social learning involves two key elements: knowledge as content (know-what) and social relations as contexts (know-who). Consumers acquire knowledge from their social contexts in two ways: by receiving information from others, and by observing others' behaviors. Lab experimental studies show that consumers are more likely to follow advice from others than to copy others' behaviors (Celen, Kariv, & Schotter, 2010). For example, consumers tend to make decisions on which movies they would like to see based on their peers' feedback on movies (Moretti, 2011). Consumers do not have sufficient financial knowledge to make informed decisions and plans for their future. Research has shown that college students' financial behaviors (e.g., saving, budgeting) are positively associated with their opportunities for acquiring advice from their close social contacts such as parents and friends (Gutter, Garrison, & Copur, 2010).

Word-of-Mouth (WOM) is a bidirectional transfer of consumer knowledge, defined as the exchange of opinions and information about specific products, services, or brands between a perceived noncommercial communicator and a receiver (Arndt, 1967; Bickart & Schindler, 2001; Bone, 1995). Consumers trust word-of-mouth as the most reliable source of product information (Hawkins, Best, & Coney, 2004; Lau & Ng 2001).

With the high rate of internet penetration, consumers are actively engaged in sharing their experiences and views about products or services with others via social media. Online WOM has attracted researchers' attention in recent years. Learning from others in online communities could increase consumer product knowledge and make consumers "more assertive and undertake reflexive behaviors" (Huang & Li, 2007, p. 493). This effect may apply to consumers differently depending on factors such as consumer involvement and the categories of targeted products (Chevalier & Mayzlin, 2006; Duan, Gu, & Whinston, 2008; Godes & Mayzlin, 2004; Liu, 2006; Park, Lee, & Han, 2007; Riegner, 2007).

In online health-related communities, Jayanti and Singh (2009) found that consumers' social learning through online communication promoted empowered decision-making. Jayanti and Singh (2009) noted the importance of pragmatic learning and defined pragmatic learning theory as "an inquiry-action framework for distributed consumer learning in online communities" (p. 1060). They described the consumer social learning process as a progressive inquiry-action process including four elements, (1) problem-focused experience, (2) inquiry, (3) knowledge generation (i.e., reflecting, reframing and exploring), and (4) action-focused experience. First, from self-experience, individuals identify problems, which motivate inquiries ("why" and "what if" questions). Then, inquiries motivate individuals to participate in collective learning by reflecting (i.e., interpreting experience to form assertions or beliefs), refining (i.e., reframing problems by integrating others' and self-experience), and exploring knowledge (i.e., developing hypotheses). Finally, individuals take actions based on the knowledge acquired in the learning stages.

## The Social Element of Social Learning

In a community, individual learning takes place at two levels. At the micro level, an individual learns from other individuals. At the macro level, an individual learns within a community in which individuals share group consensus, norms, values, and goals and have feelings of identification, belonging, and trust (Small & Supple, 2001). Community is an essential construct in social relations. Community of practice is a theory of social learning proposed by Wenger (1999). According to Wenger (2006), "communities of practice are groups of people who share a concern or a passion for something they do and learn how to do it better as they interact regularly" (p. 1). This definition indicates three elements of a community of practice: know-what (i.e., "a concern or a passion for something"), know-how (i.e., "learn how to do it), and know-who (i.e., "as they interact regularly"). In a community of practice, sustained relationships between group members are the bases of learning.

In marketing, researchers have examined properties of brand community, defined as "a specialized, non-geographically bound community, based on a structured set of social relations among admirers of a brand" (Muniz & O'Guinn, 1995; 2001, p. 412). Within brand communities, there are four ways in which consumers can create value: social networking, impression management, community engagement, and brand use (Muniz & O'Guinn, 2005). Social networking practices are related to establishing, enhancing, and maintaining social ties with other members in the brand community such as welcoming new members. Impression management practices focus on creating and maintaining favorable impressions of the brand, brand enthusiasts, and the brand community. Community engagement practices reinforce members' engagement in the

brand community and primarily focus on members' distinctive activities such as documenting unique experiences with the brand. Brand use practices are related to the exchange of knowledge about ways to enhance experiences in the use of the product (e.g, use of advanced functions of products) among members.

Research suggests that computer-supported collaborative learning (CSCL) includes two major components: effective discussion and social interaction. Effective discussion is defined as elaboration of new ideas, comments, or opinions (Hsi & Hoadley, 1997). Social interaction plays an important role in motivating group cohesion and belonging and thus is the first step for collaborative learning (Henri, 1992; Wegerif, 1998). Kreijns and Kirschner (2001) further developed this idea and proposed a theoretical framework for the positive association of social interaction and effective discussion in the environments of CSCL. According to this framework, members in a CSCL community first need to interact with the community, either with groups or individuals (termed social affordance), to develop their impressions of the community and other members in the community. Such impressions will guide members in initiating discussions or replying to the other members in specific discussions (termed social intention). All members and discussions are embedded in a social environment or social space in which social affordance facilitates social intentions.

From a community point of view, social learning within online forums is a collective process in which consumers as a group learn from each other's experience as they interact and develop relationships (Jayanti & Singh, 2009; Lave & Wenger, 1990). Many consumers have formed communities through the internet related to consumption interests such as brand interests (Kozinets, 1997, 1998). Kozinets defined the online (or

virtual) community of consumption as "affiliative groups whose online interactions are based upon shared enthusiasm for, and knowledge of, a specific consumption activity or related group of activities" (Kozinets, 1999, p. 254).

The membership of an online consumption community is determined by two nonindependent factors: (1) the relationship that an individual has with the consumption activity (consumption interests) and (2) the intensity of this relationship that an individual has with other members of the online consumption community (social interests; Kozinets, 1999). Based on the two factors, consumers can be classified into four types (in order from weak to strong connections to the group): tourists, minglers, devotees, and insiders (Kozinets, 2002, p. 64). Tourists have weak consumption interests and weak social ties with other members of online communities. Minglers have weak consumption interests but strong social ties with other members of online communities. Contrastingly, devotees have strong consumption interests but weak social ties with other members. Insiders have both strong consumption interests and strong social ties with other members.

Researchers and practitioners have primarily used participation behaviors as a measure of engagement in online communities (e.g., engagement with websites). Engaged users are defined as users who have high scores on variables such as frequency of site visits, number of page views, and number of posts made by users (Burke, Marlow, & Lento, 2009; Calder, Malthouse, & Schaedel, 2009; Faraj & Johnson, 2011; Glasgow et al., 2011; Messner & Eford, 2011; Pawan, Paulus, Yalcin, & Chang, 2003; Ransbotham & Kane, 2011). For example, researchers have used the average number of posts per student to measure students' engagement in online threaded discussions (Pawan et al., 2003).

Researchers have also used different participation behaviors to measure the level of engagement in online communities. For example, the number of threads initiated by one user is employed to measure the user's interaction with the community, whereas the number of posts made by one user is employed to measure the user's interaction with other specific users (Bateman, Gray, & Butler, 2011; Joyce & Kraut, 2006).

## Tie Strength and Social Learning

Social Network Theory (SNT) views relationships in terms of nodes and ties (SNT; reviewed in Scott 2000). A node is defined as an actor in the network, and a tie is defined as the connection between two nodes. The social network approach extends beyond the specific attributes of individual actors and focuses on relationships between actors over time (Breiger, Carley, & Pattison, 2003; Wasserman & Galaskiewicz, 1994). Thus, a social network approach can be helpful in understanding the dynamism of relations among consumers in online communities.

Tie strength is an important concept in social network theory (SNT) and measures the strength of relationships. Granovetter (1973) introduced the concept of tie strength defined as "a combination of the amount of time, the emotional intensity, the intimacy (mutual confiding), and the reciprocal services which characterize the tie" (p. 1361). Granovetter (1973) proposed that two types of ties develop in social relationships: weak ties and strong ties.

Strong ties develop when one person has frequent and reciprocal contacts with others whose social circles overlap (e.g., relatives, friends, neighbors). Weak ties are those shared between mere acquaintances or infrequent and nonreciprocal contacts (e.g.,

friends of friends). Tie strength plays an important role in bridging nodes within social structures. Strong ties tend to produce redundant information since the participants are likely to already have been exposed to the same information through their shared social circles. However, weak ties extend the reach of social networks and lead to the discovery and dissemination of new information (Granovetter, 1973, 1983). On the other hand, strong ties deepen the connections between members of the social network and thus are trusted sources of social and emotional support (Granovetter, 1973). Strong ties have greater motivation to provide support and are easily accessible. People in insecure positions are more likely to turn to strong ties for protection (Granovetter, 1983; Krackhardt, 1992; Pool, 1980). Strong ties improve mental health (Schaefer, Coyne, & Lazarus, 1981) and benefit job-seekers (Tahmincioglu, 2008).

Research in management has identified the effects of tie strength on knowledge management (Baer, 2012; Phelps, Heidl, & Wadhwa, 2012). Social networking creates channels for knowledge dissemination within organizations (Uzzi & Lancaster, 2003). Weak ties facilitate the distribution of new information, so they are important factors for knowledge creation and transfer (Levin & Cross, 2004). Borgatti and Cross (2003) proposed four characteristics of relationships relevant to information seeking: (1) knowing a person's expertise, (2) valuing that person's expertise, (3) being able to gain timely access to that person's expertise, and (4) perceiving that acquiring that person's expertise would not be too costly.

We develop our conceptual framework based on existing literature on the knowledge and social elements of social learning and the effects of tie strength on

learning. Specifically, we seek insights into the association between know-what and know-who of social learning in online communities.

## CONCEPTUAL FRAMEWORK

In this section, we first develop three propositions focusing on three key constructs: engagement, advanced learning, and sociality. We then conduct a preliminary study to examine the effects of the three constructs on referencing relationships between posters. Finally, based on the results of the preliminary study and related literature, we construct hypotheses for the dynamic development of reference relationships. We present the definition of key terms used in this study in Table 1.

Table 1. Definition for Key Terms

| Term | Definition |
|---|---|
| Post | A specific message made by a poster within a thread |
| Thread initiator | The poster who makes the first post within a thread |
| Responder | Any of the posters within a thread other than those thread initiators |
| New post | An incoming post into an ongoing thread |
| New poster | The poster of a new post |
| Referenced post | The post that a poster specifically mentions in the post |
| Referenced poster | The poster of a referenced post |
| New referenced post | The post referenced in a new post |
| New referenced poster | The poster of the new referenced post |
| Tie | A reference relationship between two posters |

## Propositions

### Proposition 1: Engagement

Individuals frequently participate in multiple threads within a given community. Their activities in threads beyond the focal or ongoing thread can be a signal of their engagement in the community. Engaging in activities such as learning the rules of posting and reading and responding in threaded discussions in an online community take time and effort. Individuals may build trust by engaging in such activities in online communities. People are more willing to develop relationships with individuals whom they trust (e.g., Hazan & Shaver, 1987). Trust is a key mediator of the effects of relationships on learning (Levin & Cross, 2004; Wenger, 2006).

Research suggests that a user's engagement in an online community positively predicts the user's activities and responses acquired from other users in the community. For example, the length of a poster's engagement in an online community positively predicts the number of threads read by the poster (Bateman, Gray, & Butler, 2011). The log-in frequency of a user positively predicts the log-in frequency of this user's friends (i.e., users listed in this particular user's friend list in one community; Trusov, Bodapati, & Bucklin, 2010). We propose that in ongoing threads, posters are more likely to be involved in reference relationships (referencing others or being referenced by others) when they have higher engagement in previous community activities. Our first proposition is

> **Proposition 1:** Posters' engagement in an online community may facilitate reference relationships in an ongoing thread.

## Proposition 2: Advanced Learning

Jayanti and Singh (2009) articulate the consumer social learning process as a progressive inquiry-action process including four elements, (1) problem-focused experience, (2) inquiry, (3) knowledge generation (i.e., reflecting, reframing and exploring), and (4) action-focused experience. Advanced levels of learning (inquiry and knowledge generation) may invoke critical thinking and thus encourage more interactions between posters, leading to reciprocal and frequent reference relationships between posters. Research has shown that posts containing questions and factual expertise are more likely to obtain more responses from users in online communities (Adamic, Zhang, Bakshy, & Ackerman, 2008; Joyce & Kraut, 2006).

Social learning is a cyclical process (Jayanti & Singh, 2009). Communication changes as relationships develop (Walther, 1992). According to the theory of cooperative learning (Brown & Palinscar 1989), one speaker tends to clarify, justify, and elaborate attitudes, opinions, and beliefs in a dialogue (offline). Researchers in online education have suggested that students develop and elaborate their ideas and opinions when they reference others' posts in online threaded discussions (Ahern, Peck, & Laycock, 1992; Bullen, 2007; Ingram & Hathorn, 2004; Pena-Shaff & Nicholls, 2004; Sze, 2008; Wise, Hamman, & Thorson, 2006). We argue that reference relationships established through previous posts may influence the content of following posts facilitating progress toward more advanced levels of social learning (e.g., knowledge generation). Posts involved in reference relationships (being referenced by others or referencing others) are more likely to contain content indicating advanced learning steps. Our second proposition is

**Proposition 2:** Posts that contain more advanced learning content including

inquiry and knowledge may facilitate reference relationships in an ongoing thread.

## Proposition 3: Sociality

Written social communication is an important component of online communities (Ellison, 2007; Schau, Muñiz, & Arnould, 2009). There are two major types of social messages in online communities, social functional messages (e.g., greetings and gratitude) and expressive messages of personal feelings and emotion (e.g., happy; Chen & Wang, 2009; Ingram & Hathorn, 2004; Schau, Muñiz, & Arnould, 2009). Previous work on online education highlights the positive role of social communication in online discussion groups (Chen & Wang, 2009; Ingram & Hathorn, 2004). Social messages are off-task but "important in creating an environment that supports collaboration, such as introductions…" (Ingram & Hathorn, 2004, p. 225). Research in brand community suggests that social messages help create, enhance, and maintain ties among users (Schau, Muñiz, & Arnould, 2009). We propose that an online community is a social networking environment in which participants learn from others as they develop reference relationships with others. Our third proposition is

**Proposition 3:** Posts that contain more social content may facilitate reference relationships between posters in an ongoing thread.

## Preliminary Study: Likelihood of Being Referenced

In this preliminary study, we investigate the effects of engagement, advanced learning, and sociality on the likelihood of a post being referenced in an ongoing thread

in a diabetes online community. The goal of this preliminary study is to provide empirical evidence to develop hypotheses about the dynamic development of reference relationships for the main study.

## Methodology

In this study we use mixed methods to address the research questions. First, we use qualitative analysis to explore data collected from an online forum related to diabetes. We develop and implement a coding scheme allowing us to create measures of patterns of communication and content of posts in the threads analyzed. Second, we apply quantitative analysis to the data collected in the first step in order to determine the association of reference relationships in an ongoing thread and posters' engagement in community activities and the content of communication in the ongoing thread.

## Study Setting

According to The Centers for Disease Control and Prevention (CDC), diabetes is one of the major health concerns facing US consumers: in 2011, prediabetes affected 35% of adults age 20 and older and half of Americans age 65 and older; diabetes affected 8.3% of all Americans and 11.3% of adults age 20 and older. CDC estimates that as many as one in three U.S. adults could have diabetes by 2050 if current trends continue. Diabetes can lead to serious health complications and premature death, but people with diabetes can control the disease by taking steps such as maintaining a healthy diet. Because of the growing population with diabetes and the significance of self-care in managing the disease, we selected one diabetes forum (http://www.diabetesforum.com)

as the online community for this research. At the time the data for this study were collected, the forum had 25,000 members.

## Data Collection

We randomly downloaded 15 threads with a total of 218 posts, with each thread containing at least four posts. The total number of posts in one thread ranges from five to 48 posts. A variety of information useful in understanding the context of the forum was available on the website. The web page of the forum displayed the titles of all the discussion threads. Clicking on the title of each thread showed the thread page and the data used for this study, including the complete text of each post, the time each post was posted, and the name (online identity) of individual posters.

A hyperlink from each poster's name to each poster's profile page was also available. The profile page showed the statistical data of the poster's communication history in the forum, including the poster's average number of posts per day, the total number of "likes" that the poster received from other posters, and the total number of "likes" that the poster gave to other posters. The Like button is a feature in the forum that allows users to express their interest in a specific post by clicking on the "Like" button at the end of one post.

## Coding and Measures

We read each post and then coded the content by hand. We diagramed each message (one post) into statements. A statement is a complete sentence or a complete idea within a sentence. Based on previous work on online learning (Chen & Wang, 2009;

Ingram & Hathorn, 2004; Jayanti & Singh, 2009), we categorized the content of posts into four categories: experience, inquiry, knowledge, and sociality (see Table 2).

For each post, we recorded the order in which posts were made in the thread and the name of the poster. If one post referenced another post, we recorded the order of the referenced post and the name of the referenced poster. The references are clear if the posters use the "Reply with quotes" command. For posts without "quotes," we coded the references by identifying terms such as "your message…" or "Yes, I agree." For each post, we coded the reference relationship as "1" if one post was referenced by other posts or "0" otherwise.

Granovetter (1983) proposed four dimensions of tie strength: amount of time, intimacy (mutual disclosure), emotional intensity, and reciprocal services. Subsequent research has identified additional dimensions, including communication reciprocity (Friedkin, 1980), shared relationships (Shi, Adamic, & Strauss, 2007), recency of communication (Lin, Dayton, & Greenwald, 1978), and interaction frequency (Gilbert, Karahalios, & Sandvig, 2008). We extended the work on tie strength in the context of online threaded discussions. We categorize tie strength along two dimensions: (1) reciprocity (two posters reference posts of each other) and (2) frequency (how often one poster references posts by the other poster). For any two posters with at least one poster making at least one post to the other, we coded the strength of the ties between those posters into three types from high strength ties to weak ties (type I, II, III, respectively) based on reciprocity and frequency as shown in Table 3. The strongest ties feature reciprocal and frequent occurrence of referencing of posts between two posters. The weakest ties feature nonreciprocal referencing between posts by two posters. We define

Table 2. Coding Scheme for Content

| Category | Description | Example |
|---|---|---|
| Experience | Statements related to own or others' experiences | "I inject 30 units of Lantus each morning " |
| Inquiry | Questions | "Are you using medicine?" |
| Knowledge | Statements of opinions or ideas, factual information | "Lantus is a very slow and long acting insulin" |
| Sociality | Supportive statements,and statements related to personal feelings and desires. | "that is awesome" |

Table 3. Types of Strength of Ties

| Tie strength | Dimensions of reference relationship | |
|---|---|---|
| | Reciprocity (yes/no) | Frequency (multiple/one-time) |
| Type I (strongest ties) | Yes | Multiple times |
| Type II | Yes | One-time |
| Type III (weakest ties) | No | One-time or multiple times |

the reference relationship for posts that do not reference any specific posts as "reference to the group."

From each poster's profile page, we acquired four types of historical activity data: (1) the number of posts made by the poster as of the date the study data were collected, (2) the number of likes received by the poster, (3) the number of likes given by the poster, and (4) the number of threads initiated by the poster as of the date the posts were made. Based on previous work on online communities (e.g., Bateman, Gray, & Butler, 2011), we have two measurements for posters' engagement in the community, including engagement with individual posters in the community and engagement with the community as a whole. In online communities, making posts and receiving and giving likes indicate posters' engagement with other posters; the initiation of threads indicates posters' engagement with the community.

We first conducted analysis with one post as the unit of analysis. The reference relationship of a post in an ongoing thread is measured by whether or not the post is referenced by other posts. Each poster's engagement in the community (either with the community or with individuals) is measured by the poster's historical activity data. The content of posts is measured by the number of statements in each category of content (e.g., sociality) in the post.

We then conducted analysis with one thread as the unit of analysis. The strength of ties established in an ongoing thread is measured by the percentage of posts in the thread at each level of tie strength (type I, II, III, and reference to the group). The engagement of posters in each thread is measured by averaging the historical activity data of the posters in the thread. The content of a thread is measured by the percentage of

statements in the thread in each category of content.

## Analysis

We first used the social network software UCINET to generate a graph of patterns of ties for each of the 15 threads. In these graphs, nodes represent unique posters within the group, and lines with arrows represent the ties between posters and between posters and the group. The width of the lines represents the strength of ties. The graphs (Figure 3) show that the patterns of ties within threads vary widely in the number of nodes and ties, and the strength of ties. Threads 1 to 7 show networks constructed with weak ties, whereas threads 8 to 15 show networks constructed with strong ties.

We undertook correlation analysis, factor analysis, and logistic regression analysis to test the propositions using SPSS. We first conducted factor analysis with varimax rotation on the variables used to measure posters' engagement with individual posters (number of posts, number of likes received, and number of likes given). These variables all loaded together with regard to analyses at both the level of posts and the level of threads (see factor loadings in Table 4). SPSS provides regression based factor scores for each subject. We acquired the factor score from SPSS and used it in further analyses of ties.

We analyzed reference relationships at two levels: the post level and the thread level. At the post level, we focused on the establishment of each reference relationship. When a post is referenced by another post, social learning is continued between the posters of the two posts. Further, within the thread, two individual posters may strengthen their ties by continually making posts to each other, forming a type I tie (reciprocal and

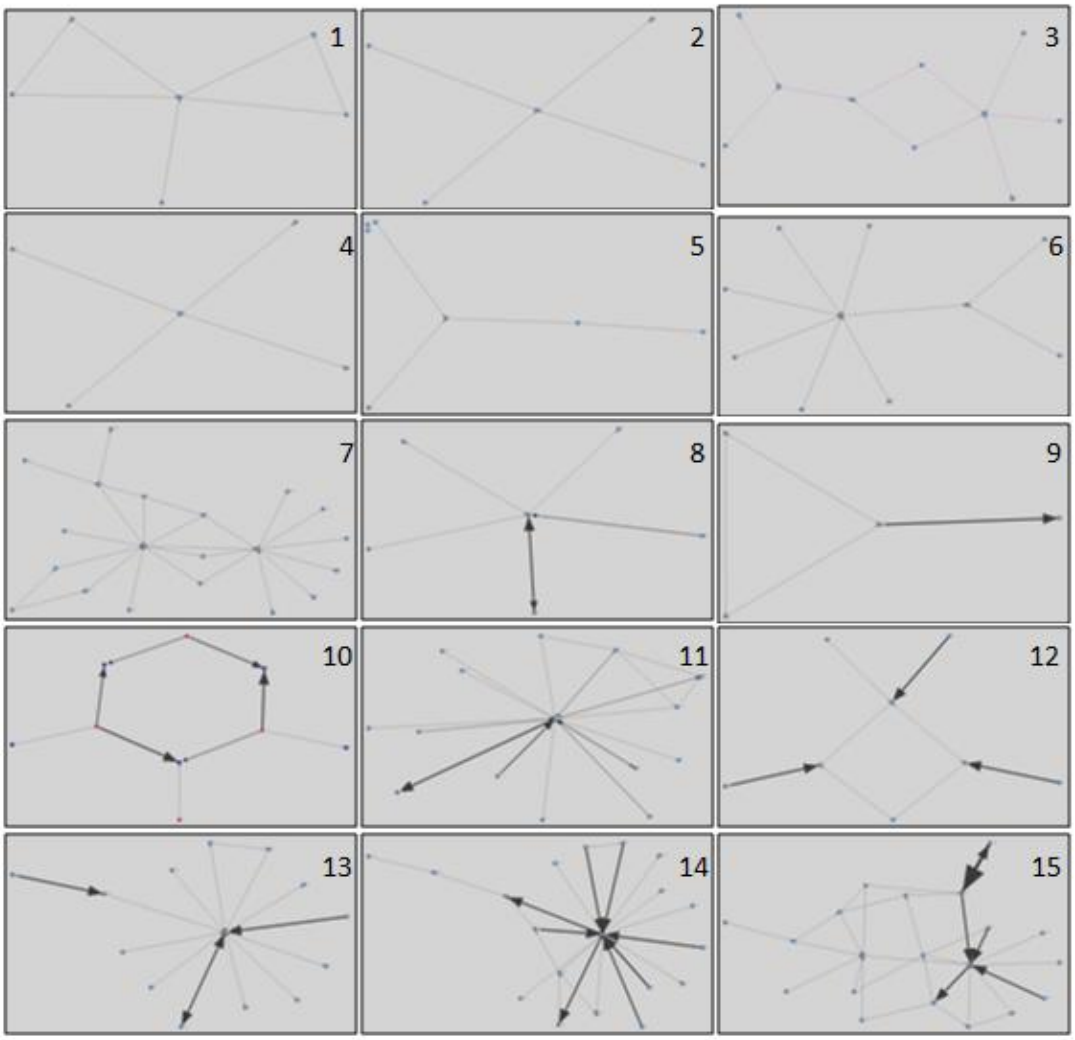


Figure 3. Graphs for Social Networks Representing 15 Threads

Table 4. Factor Analysis Results

| Variable | Factor Loading | |
|---|---|---|
| | Post as the unit of analysis | Thread as the unit of analysis |
| Number of posts per day | .833 | .832 |
| Number of likes received | .871 | .967 |
| Number of likes given | .918 | .968 |

frequent). At the thread level, we focus on the strength of ties by aggregating the ties within the thread with the same type of tie strength. Our study aims to identify the factors associated with the establishment of each level of tie strength as well as the strength of aggregate ties within threads.

For the analysis at the post level, we conducted binary logistic regression analysis to investigate the effects of posters' engagement and the content of posts on the reference relationships in the ongoing thread. Among the 218 posts, 116 posts (52%) were coded as "0" (not referenced by other posts) and 103 posts (48%) were coded as "1" (referenced by other posts). For the analysis at the thread level, we then performed Pearson's correlation analysis to explore the relationship between the strength of ties and posters' engagement and the content of posts.

## Results

Overall, the results provide substantial evidence that tie strength demonstrated by reference relationships in an ongoing thread is significantly associated with posters' engagement in the community and the content of posts in the ongoing thread. We present the analysis results in this section.

### Reference Relationship in the Ongoing Thread

Descriptive statistics for all of the variables included in the analysis are presented in Table 5. Within the 218 posts, there are a total of 856 statements indicating experience, 661 statements indicating knowledge, 87 statements indicating inquiry, and 206 statements indicating sociality. The categories of experience and knowledge are the top

Table 5. Descriptive Statistics for Posts

| Measure | Minimum | Maximum | Mean | *Std.* Deviation |
|---|---|---|---|---|
| Post is referenced | 0 | 1 | 0.468 | 0.500 |
| Engagement | | | | |
| Number of posts per day | 0 | 11 | 2.760 | 3.049 |
| Number of previous threads | 0 | 268 | 35.530 | 51.162 |
| Number of likes received | 0 | 6264 | 416.771 | 690.545 |
| Number of likes given | 0 | 5412 | 471.355 | 1023.591 |
| Content | | | | |
| Number of statements indicating experience | 0 | 21 | 3.339 | 2.965 |
| Number of statements indicating inquiry | 0 | 4 | 0.151 | 0.490 |
| Number of statements indicating knowledge | 0 | 22 | 3.032 | 3.737 |
| Number of statements indicating sociability | 0 | 28 | 0.817 | 2.433 |

two categories of content of the threaded posts in the community.

The Hosmer and Lemeshow Test for goodness of fit shows a nonsignificant chi-square indicating that the data fitted the model well ($p = 0.145$). Table 6 shows the logistic regression coefficients, Wald tests, and odds ratios for each of the predictors. The coefficient for engagement with individual posters ($B = -0.527$) is negative and significant at the $p = 0.05$ level, suggesting that as engagement with individual posters increases, the likelihood of being referenced will decrease. The coefficient for the number of statements indicating inquiry ($B = 1.468$) is positive and significant at the $p = 0.01$ level, suggesting that as the number of statements indicating inquiry increases, the likelihood of being referenced will increase.

The coefficients of two predictors are positive and significant at the $p = 0.1$ level. As engagement with the community ($B = 0.008$) increases, the likelihood of being referenced will increase. As the number of statements indicating sociality ($B = 0.268$) increases, the likelihood of being referenced will increase.

Tie Strength in the Ongoing Thread

In the analysis of tie strength in an ongoing thread, variables of interest (engagement and content of posts) are aggregated across all of the posters in the thread. The descriptive statistics are shown in Table 7. Results of Pearson correlation analysis show that engagement with the community was positively correlated with the percentage of posts with type I ties, high frequency and reciprocal postings between two posters ($R^2 = 0.543$, $p = 0.036$), and negatively correlated with the percentage of posts in the ongoing thread referencing the group ($R^2 = -0.594$, $p = 0.02$). These results show that the stronger

Table 6. Logistic Regression Results

| Predictor | *B* | *Wald* | *P* | *Odds* |
|---|---|---|---|---|
| Engagement | | | | |
| Engagement with the community | 0.008 | 2.87 | 0.09 | 1.008 |
| Engagement with individual posters | -0.527 | 5.372 | 0.02 | 0.59 |
| Content | | | | |
| Experience | -0.081 | 2.211 | 0.137 | 0.923 |
| Inquiry | 1.468 | 8.492 | 0.004 | 4.341 |
| Knowledge | -0.022 | 0.293 | 0.589 | 0.978 |
| Sociality | 0.268 | 3.271 | 0.071 | 1.307 |
| Constant | -0.515 | 3.578 | 0.059 | 0.597 |

Table 7. Descriptive Statistics for Individual Threads

| Measures | Minimum | Maximum | Mean | *Std.* Deviation |
|---|---|---|---|---|
| Tie Strength | | | | |
| Percentage of posts with type I tie | 0.000 | 0.857 | 0.239 | 0.286 |
| Percentage of posts with type II tie | 0.000 | 0.667 | 0.176 | 0.228 |
| Percentage of posts with type III tie | 0.071 | 0.750 | 0.337 | 0.220 |
| Percentage of posts referencing to the group | 0.048 | 0.714 | 0.248 | 0.171 |
| Engagement | | | | |
| Average number of previous threads | 1.429 | 65.333 | 27.930 | 13.877 |
| Average number of posts per day | 0.281 | 4.352 | 2.144 | 1.127 |
| Average number of likes received | 0.000 | 743.111 | 321.509 | 250.823 |
| Average number of likes given | 0.000 | 861.889 | 374.020 | 290.452 |
| Content | | | | |
| Percentage of statements indicating Experience | 0.085 | 0.708 | 0.405 | 0.201 |
| Percentage of statements indicating inquiry | 0.000 | 0.167 | 0.063 | 0.044 |
| Percentage of statements indicating knowledge | 0.167 | 0.745 | 0.383 | 0.162 |
| Percentage of statements indicating sociality | 0.000 | 0.339 | 0.112 | 0.107 |

posters' engagement with the community, the higher percentage of posts in the ongoing thread with the strongest ties. However, the weaker posters' engagement with the community, the higher the percentage of posts in the ongoing thread referencing the group.

Posters' engagement with individual posters was positively correlated with the percentage of posts with type III ties, nonreciprocal posting between two posters ($R^2$ = 0.517, $p$ = 0.048). Threads involving posters who have strong engagement with individual posters had a higher percentage of posts with low strength ties. The participation of posters with strong engagement with individual posters in threads is associated with less reciprocity among posters.

The percentage of statements in the ongoing thread indicating experience was negatively correlated with the percentage of posts with type I ties ($R^2$ = -0.562, $p$ = 0.029) and positively correlated with the percentage of posts referencing the group ($R^2$ = -0.755, $p$ = 0.001). The percentage of statements in the ongoing thread indicating knowledge is negatively correlated with the percentage of posts referencing the group ($R^2$ = -0.521, $p$ = 0.046). The percentage of statements indicating inquiry is negatively correlated with the percentage of posts in the ongoing thread with type III ties ($R^2$ = -0.461, $p$ = 0.084). The percentage of statements indicating sociality is positively correlated with the percentage of posts in the ongoing thread with type I ties ($R^2$ = 0.777, $p$ = 0.001) and negatively correlated with the percentage of posts referencing the group ($R^2$ = -0.62, $p$ = 0.014). Results of Pearson's correlation analysis are shown in Table 8.

The results of the analysis of correlation of tie strength and content of posts indicates that within one thread, (1) the higher the percentage of posts with high strength

Table 8. Results of Pearson's Correlation Analysis

| Measures | Type I tie | | Type II tie | |
|---|---|---|---|---|
| | *R* square | *p* | *R* square | *p* |
| Engagement | | | | |
| Engagement in interaction with the community | 0.543** | 0.036 | -0.297 | 0.282 |
| Engagement in interaction with individual posters | -0.226 | 0.417 | -0.065 | 0.819 |
| Content | | | | |
| Experience | -0.562** | 0.029 | 0.036 | 0.898 |
| Knowledge | 0.161 | 0.566 | -0.036 | 0.898 |
| Inquiry | -0.156 | 0.580 | 0.347 | 0.205 |
| Sociality | 0.777*** | 0.001 | -0.143 | 0.611 |
| Measures | Type III tie | | Ref. the group | |
| | *R* square | *p* | *R* square | *p* |
| Engagement | | | | |
| Engagement in interaction with the community | 0.014 | 0.960 | -.594** | 0.020 |
| Engagement in interaction with individual posters | 0.517** | 0.048 | -0.181 | 0.518 |
| Content | | | | |
| Experience | 0.012 | 0.967 | 0.755*** | 0.001 |
| Knowledge | 0.304 | 0.271 | -0.521** | 0.046 |
| Inquiry | -0.461* | 0.084 | 0.431 | 0.109 |
| Sociality | -0.345 | 0.207 | -0.620** | 0.014 |

Notes: *$0.05 \leq P < 0.1$, **$0.01 \leq P < 0.05$, ***$0.001 \leq P < 0.01$

ties, the higher the percentage of content indicating sociality and the lower the percentage of content indicating experience; (2) the higher percentage of posts with low strength ties, the lower the percentage of content indicating inquiry; and (3) the higher percentage of posts referencing the group, the higher the percentage of content indicating experience and the lower the percentage of content indicating knowledge and sociality.

## Discussion

Research in consumer social learning has been increasing in recent years. However, there has been limited research about the mechanisms by which relationships among consumers develop during the social learning process. In this study, in the context of an online community, we used the reference relationships of posts to assess tie strength in ongoing threads. Tie strength is measured by reference relationships along the two dimensions of reciprocity and frequency. Our propositions address two questions: (1) the effects of the posters' engagement in community activities and the content of the posts on the possibility of the post being referenced in other posts and (2) the correlation between tie strength in the thread and posters' engagement in community activities and the content of posts. Our analysis demonstrates consistent results: tie strength in ongoing threads was significantly associated with posters' engagement in community activities and the content of posts.

The results provide strong support for the association between posters' engagement in the online community and tie strength in ongoing threads. This suggests that posters' engagement in the community plays an important role in the consumer social learning process. Consumers develop engagement through two types of community

activities: interaction with the community and interaction with individual members of the community. Engagement with the community is positively correlated with the percentage of posts with high strength ties (reciprocal and frequent interaction between posters) and the possibility of one post being referenced by other posts. Engagement with individual members of the online community is positively correlated with the percentage of posts with low strength ties between individual posters (nonreciprocal interaction with individual posters) and negatively correlated with the possibility of a post being referenced by other posts.

The different effects of engagement with the community and with individual members in the ongoing thread are notable. Engagement with the community is established by initiating threads, whereas engagement with individuals results from such activities as giving a "like" to an individual member's post. We argue that thread initiators are driven by either or both of two factors: (1) their relationship with the community (e.g., having trust in the community) and (2) their interest in the consumption activities (e.g., seeking answers to questions related to consumption activities). However, engagement with individual posters appears to be related to a single factor: interest in the consumption activities mentioned in the post made by an individual member. Posters with different levels of engagement with community and engagement with individuals demonstrated different behaviors in reference relationships.

Posters' engagement and the content of posts did not show significant effects on the midstrength ties (reciprocal and one-time communication between posters). A larger sample of threads may be needed to accurately assess the relationship between midstrength ties and engagement and content of posts.

The results provide strong support for the association between tie strength and content of posts in ongoing threads indicating progressive steps in the social learning process. The percentage of posts with high strength ties (reciprocal and frequent communication between two posters) was positively associated with the percentage of statements indicating sociality and negatively associated with the percentage of statements indicating experience. Contrastingly, the percentage of posts referencing the group was positively associated with the percentage of statements indicating experience and negatively associated with the percentage of statements indicating knowledge and sociality. When there are more posts in an ongoing thread with high strength ties, the content of posts in the thread may focus more on sociality and less on lower levels in the learning process such as experience. Likewise, when there are more posts referencing the group, the content of posts in the thread may focus more on lower levels in the social learning process and less on sociality and higher levels of social learning such as knowledge.

The percentage of statements indicating inquiry is negatively associated with the percentage of posts with low strength ties (nonreciprocal communication between posters) and positively associated with the possibility of posts being referenced in other posts. This shows that posters who ask questions in their posts are more likely to get "quoting" response from other posters. The percentage of statements indicating sociality is also positively associated with the possibility of posts being referenced by other posts, suggesting that posters who demonstrate more sociability in their posts are more likely to get "quoting" responses from other posters.

## Hypotheses

The logistic regression analysis in the preliminary study predicts the likelihood of a post being referenced by other posts in one thread by investigating this poster's engagement in the community and the content of the post in the thread. A reference relationship includes two parties: the referenced posters and the poster who reference**s** others (excluding the thread initiator). The logistic regression analysis focuses on one of the two parties: the posters who are referenced by other posters. To get a broad picture of reference relationships, we conduct Pearson correlation analysis of tie strength measured by the frequency and reciprocity of reference relationships and engagement of posters in the on-going thread and the content of posts in one thread.

Based on Kozinets' (2002) theory of online community of consumption and the results from the preliminary study, we suggest that there are two types of posters who are more likely to reference other posters. One type of poster is interested in relationships with both consumption activities and other members in the community and so is highly engaged with the community. The other type of posters is only interested in relationships with consumption activities and therefore is highly engaged with individual posters. In the main study, we take a dynamic view to look at the patterns of referencing others by new posters in an ongoing thread. Specifically, we investigate the new posters' engagement in the ongoing thread and across threads in the broader community over time, the content of the new posts in the ongoing thread, and the ties between the new posters and the new referenced posters. This study considers the development of relationships over time by focusing primarily on analysis at the post level.

In an ongoing thread, a new post can reference the thread initiator, a responder, or

the group. Common sense suggests that a responder will make a new post in response to the thread initiator or the group by providing information relevant to the question proposed by the thread initiator. The more interesting reference relationships are those between new posters and responders in which new posters and responders are involved in "new" discussions relevant to the question advanced by the thread initiator. We view the decision by a new poster of who to reference as an initial decision. If the decision is to reference a responder (rather than the group), then a second (contingent) decision is which responder to reference, a new friend or an old friend. A new friend is defined as a poster by whom the new poster had not been referenced directly in the ongoing thread. An old friend is defined as a poster by whom the new poster had been referenced directly in the ongoing thread. This study focuses on the reference relationships between new posters and responders by investigating the factors that influence the new posters' decisions about who to reference.

## Dynamic Development of Posters' Engagement

Posters build trust of others by developing their relationships with other posters across threads in the community. Our preliminary study shows that posters are more likely to reference a specific responder when the poster has higher engagement with the community, individuals, or both. We argue that a new poster's engagement in the community will be positively associated with the possibility of referencing a responder in the ongoing thread. New posters may be engaged both in the ongoing thread and in the community.

Research suggests that as participants spend more time engaging in a discussion

(e.g., responding to others) in an online community, they are more likely to contribute to knowledge collaboration (Lakhani & Hippel 2003; Rafaeli & Ariel, 2008). We argue that a poster's engagement over time in the ongoing thread indicates this poster's interest in the continued discussion in the thread. The poster with enduring engagement will be more likely to pay attention to messages posted by individual responders and thus reference the posts by individual responders. The early posters in the thread may focus more on the thread initiator, whereas the middle and latter posters may focus more on the topics discussed in which they take personal interest. Engaging in a thread longer also provides the poster more opportunities to build trust with the group and members in the group. Therefore, our first hypothesis is

**H1:** A new poster will reference a specific responder,

**a:** when the new poster has higher engagement in the community;

**b:** when the new poster has higher engagement in the ongoing thread.

In an online learning community, students tend to initiate a discussion when they have formed impressions of the community and members in the community. These impressions are formed through engagement in online activities in the community (Kreijns & Kirschner, 2001). Referencing to a new friend indicates the new poster's intention to initiate a new conversation with a responder. We expect this will be more likely when the new poster has higher engagement with the community and with the ongoing thread. Therefore, our second hypothesis is

**H2:** A new poster will reference a new friend,

**a.** when the new poster has higher engagement in the community;

**b.** when the new poster has higher engagement in the ongoing thread.

## Dynamic Development of Sociality and Advanced Learning

Our preliminary study suggests that when posters are engaged in referencing, they are more likely to demonstrate sociality and advanced learning steps including inquiry and knowledge. Referencing a specific responder shows that a poster is interested in not only the general topics under discussion in the ongoing thread, but also the specific topics raised by a responder. Based on the theory of reciprocity (Fehr & Gächter, 2000), when replying to an old friend, the poster is more likely to use social messages to demonstrate friendliness. Further, when responding to an old friend, the poster is involved in a reciprocal dialogue with the other poster. This suggests a way of collaborative learning in which posters are exchanging more other-focused knowledge, but less self-focused experience. So, our third hypothesis is

**H3:** Posts that reference a responder

**a**. include more social statements;

**b.** include more statements indicating advanced learning steps including inquiry and knowledge;

**c.** include less experience-related statements.

In an online community, old members create ties with new members by sending social messages to the new members (Muniz & O'Guinn, 2005). To create a tie with a new friend, the poster sends social messages to the new friend. In cognitive science, occasions in which more than one person is engaging in a conversation about an object are defined as joint attention (e.g., Brooks & Meltzoff, 2005; Morales et al., 2000). Initiating joint attention (IJT) occurs when a person directs another person's attention to the object. When the second person follows the conversation, this constitutes responding

to joint attention (RJT). Previous work suggests that initiating joint attention requires greater cognitive control processes than does responding to joint attention (Redcay, Kleiner, & Saxe, 2012). Referencing a new friend, by whom the new poster has not been referenced in the ongoing thread, is a way to initiate a new conversation. By engaging in this new conversation, the new poster can direct others' attention to specific topics and sadd new ideas or opinions to the discussion in the ongoing thread. Therefore, our fourth hypothesis is

**H4:** Posts that reference a new friend

**a**. include more social statements;

**b.** include more statements indicating advanced learning steps including inquiry and knowledge;

**c.** include less experience-related statements.

## Dynamic Development of Tie

In our preliminary study, we define a tie as the accumulation of reference relationships between two posters over time in an ongoing thread. In our main study, we expand our view of ties from a single type (direct ties) to two types (direct and indirect ties) and from the ongoing thread to other threads in the community.

Transitivity assumes that if a chooses b as a friend and b chooses c as a friend, a will choose c as a friend. In this relationship, b is a shared contact connecting a and c. Transitivity plays an important role in distributing knowledge and strengthening ties in social networks (Brown & Reingen, 1987; Feld, 1981; Louch, 2000; Shi, Adamic, & Strauss, 2007; White & Houseman, 2002). In our study, transitivity can be a thread or

thread initiator. In a thread, both a and b make posts, but neither of them references the other. The thread or thread initiator is a bridge linking a and b. We define a direct tie as the accumulation of reference relationships between two posters in a thread. We define a shared tie as a relationship in which two posters make posts but do not have any referencing relationships with each other in a thread.

Ties among two posters may be built over time in any thread in the community. Ties between posters established previously in other threads may influence the reference relationship of a new post in the ongoing thread. These prior ties may establish trust between the two posters (Shi, Adamic, & Strauss, 2007). People are more willing to acquire and disseminate information with their trusted contacts. When a new poster has ties with a new friend in other threads in the community, trust may have been established facilitating reference to a new friend, that is, to a poster with whom no previous references in the ongoing thread have been made

We use an example to illustrate the developments of ties between two posters across threads in the community (see Figure 4). Posters a and b may establish a tie (either a direct or shared tie) in one thread A. In the subsequent thread B, poster a may directly reference poster b who has not directly referenced a in the current thread B. We define poster b as a new friend of poster a in the current thread B. We argue that it is the familiarity or the trust built between the two posters, a and b, in the previous thread A that leads poster a to initiate a conversation with poster b in the ongoing thread B.

The strength of ties in previous threads in the community can be measured along five dimensions: coexistence, directness, recency, reciprocity, and frequency. Coexistence is defined by whether the two posters (a and b) made posts in at least one

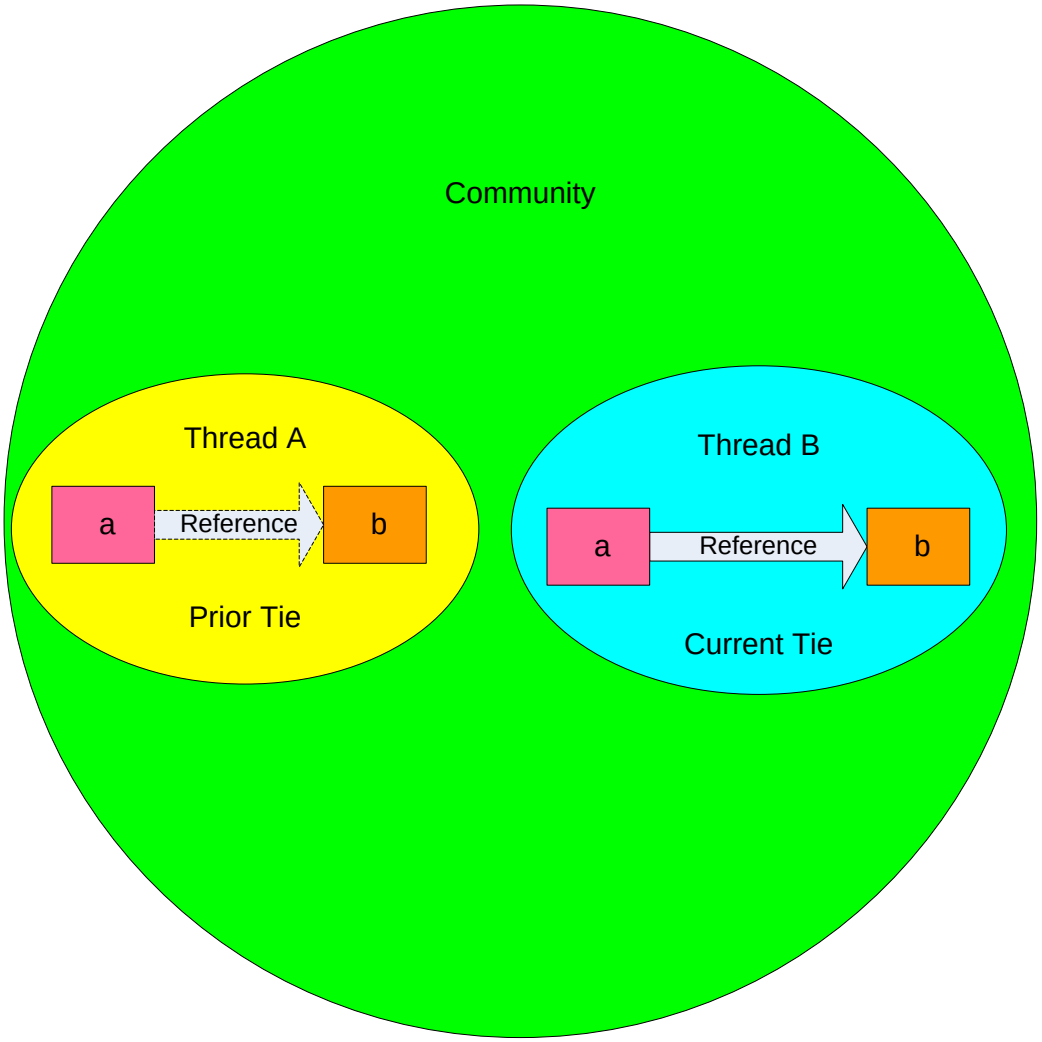


Figure 4. An Example of Ties in One Other Thread and the Ongoing Thread

other thread. Recency is defined as the length of time between when the most recent post was made between the two posters in one other thread and when the new post was made between them in the ongoing thread. Directness is defined as whether one of the two posters directly referenced the other in at least one other thread. Reciprocity and frequency are two dimensions of the strength of direct ties. Reciprocity is defined as whether the two posters referenced each other in at least one other thread. Frequency is defined as how frequent (one time or multiple times) the two posters referenced each other in one other thread.

Strong ties lead to familiarity and trust between posters. We expect that a new poster will be more likely to reference a new friend with whom they have previously established a strong tie in other threads in the community. We propose that a strong tie (i.e., shared, direct, reciprocal, frequent, and recent ties) between two posters in one thread is positively correlated with one of the two posters' initiation of a conversation with the other poster in a subsequent thread.

Therefore, our fourth hypothesis is

**H4:** A new poster who references a new friend in the ongoing thread has stronger ties with the referenced poster in other threads in the community than does a new poster who references an old friend.

# MAIN STUDY

## Methods

In this study we used mixed methods to address the research questions. First, we use qualitative analysis to explore data collected from the online diabetes forum that we investigated in the preliminary study. We develop and implement a coding scheme allowing us to create measures of patterns of reference and content of posts in the threads analyzed. Secondly, we apply quantitative analysis to test our hypotheses.

The main study aims to investigate the factors that influence the reference relationship over time in the ongoing thread. Therefore, for this study, we collect threads each containing at least 15 posts. We use six threads from the preliminary study and select the four longest threads from the same forum, those that were the longest in terms of the number of posts. All the threads are officially closed. Our sample included 10 threads, 451 posts, and 209 unique posters. By the date when all the data were collected, there were a total of 104 threads that met our requirements (officially closed and contained more than 15 posts). We collected data on posts, posters, and reference relationships between posters for each thread.

### Coding and Measures

We coded the content of posts using Yoshikoder, a cross-platform multilingual content analysis program developed as part of the Identity Project at Harvard's

Weatherhead Center for International Affairs. Yoshikoder allows users to load text documents, construct and apply content analysis dictionaries, and perform basic content analysis such as word counts and highlighting the words in each category. A dictionary is the collection of words that defines a specific category. We constructed a dictionary containing the four categories of content: experience, knowledge, sociality, and inquiry. We coded each thread by taking three steps: (1) first, loading the text document (.txt format) of each thread into Yoshikoder, (2) second, highlighting the words in one category, and (3) finally, coding each statement into a single category of content.

We constructed the coding dictionary based on Linguistic Inquiry and Word Count (LIWC) dictionary, which was psychometrically developed and validated by social psychologist James Pennebaker and his colleagues at the University of Texas at Austin. Language use reflects individuals' social and psychological processes such as cognitive processes, personality, and social relationships (Newman, Pennebaker, Berry, & Richards 2003; Pennebaker, Mayne, & Francis, 1997; Pennebaker & Graybeal, 2001; Pennebaker, Mehl, & Niederhoffer, 2003; Tausczik & Pennebaker, 2010). The LIWC dictionary has been widely used to analyze text messages in both online and offline environments (Arguello, 2006; Newman, Groom, Handelman, & Pennebaker, 2008; Slatcher, Vazire, & Pennebaker, 2008; Tumasjan, Sprenger, Sandner, & Welpe, 2010).

We used the same coding rules in the main study as we used in the preliminary study. We defined inquiry as statements indicating questions. Any statement ending with a question mark was coded as an inquiry. In the dictionary, we had a question mark "?" as an entry for the category of inquiry. We defined sociality as statements related to personal feelings. We imported the affect category (including 915 words) from the LIWC

dictionary (e.g., “happy,” “sad”) and added a category of “greeting and gratitude” including entries such as “hello” and “thanks.”

We defined knowledge as statements of opinions, comments, or ideas and experience as statements of own or others' experiences. Coding involved two steps. First, we imported the categories of personal pronouns, impersonal pronouns (e.g., “it,” “those”), past and present tense verbs, and articles from the LIWC dictionary. The category of personal pronouns include five categories: first-person singular (e.g., “I,” “mine”), first-person plural (e.g., “we,” “us”), second-person (e.g., “you,” “your”), third-person singular (e.g., “she”), and third-person plural (e.g., “they”). The category of article includes three words: “a,” “an,” and “the.” Next, we imported the category of cognitive process, which includes eight subcategories: insight (e.g., “think,” “know”), causation (e.g., “because”), discrepancy (e.g., “should”), tentative (e.g., “maybe”), certainty (e.g., “sure”), inhibition (e.g., “block”), inclusive (e.g., “include”), and exclusive (e.g., “but”). Personal pronouns reflect where people are focusing (Slatcher et al., 2008). First-person plural pronouns indicate group identity and shared values and motivation, whereas second-person, impersonal pronouns, and articles indicate other-focused attention (Slatcher et al., 2008; Tausczik & Pennebaker, 2010).

We coded a statement as a “knowledge” statement if the statement included a non-I/my related pronouns (a first-person plural pronoun, second-person pronoun, third-person pronoun, impersonal pronoun), an article, or a noun as part of the main object or subject and any word in the categories of cognitive process. We coded a statement as an “experience” statement if the statement had an I/my related pronoun (a first-person single pronoun) as the main object or object and any word in the categories of past and present

tense verbs excluding the verb to be in different forms (e.g., is, was). We present the categories that we imported from the LIWC dictionary in Table 9 and the coding scheme for content of posts in Table 10.

We coded the content of posts in each thread following these steps. Yoshikoder has a highlight function through which the words in a specific category are highlighted in yellow. We first imported a text (.txt) file containing the posts in one thread and the dictionary to Yoshikoder. We then highlighted the words or punctuations ("?") in each category and coded each statement into a single category. Words were first coded into the inquiry category. After we finished coding for that category, only the statements left uncoded were included in the next coding step, that for experience. After that, those remaining uncoded were coded into the category of knowledge, and finally, those remaining we coded into the category of sociality. Each statement was coded into only one category.[1] One word in the dictionary may be in multiple categories. For example, "think" is in the category of knowledge and the category of present verb. We ordered the coding process to avoid coding one statement into more than one category.

In the main study, we coded a total of 431 posts. Among them, we had 169 posts that were collected from the preliminary study and coded both by hand and computer. We conducted a Pearson correlation analysis between hand coded and computer coded posts. Results showed high correlations between hand-coded and computer-coded data: 94% (sociality), 96% (knowledge), 97% (experience), and 98% (inquiry). Besides the content of new posts, we also measured new posters' engagement in the community and in the ongoing thread, reference relationships in the ongoing thread, and the strength of ties

[1] If a statement is categorized into the category of experience, it was not categorized into the category of sociality.

Table 9. Categories from the LIWC Dictionary

| Category | Examples | Words In Category |
|---|---|---|
| Total pronouns | | 116 |
| Personal pronouns | I, them, her | 70 |
| 1st person singular | I, me, mine | 12 |
| 1st person plural | We, us, our | 12 |
| 2nd person | You, your, thou | 20 |
| 3rd person singular | She, her, him | 17 |
| 3rd person plural | They, their, they'd | 10 |
| Impersonal pronouns | It, its, those | 46 |
| Articles | A, an, the | 3 |
| Verbs | | 314 |
| Past tense | Went, ran | 145 |
| Present tense | Hear, take | 169 |
| Affective processes | | 915 |
| Positive emotion | Love, nice, sweet | 406 |
| Negative emotion | Hurt, ugly, nasty | 499 |
| Anxiety | Worried, fearful | 91 |
| Anger | Hate, kill, annoyed | 184 |
| Sadness | Crying, grief, sad | 101 |
| Cognitive processes | | 730 |
| Insight | think, know | 195 |
| Causation | because, effect | 108 |
| Discrepancy | should, would | 76 |
| Tentative | maybe, perhaps | 155 |
| Certainty | always, never | 83 |
| Inhibition | block, constrain | 111 |
| Inclusive | And, with, include | 18 |
| Exclusive | But, without | 17 |

Table 10. Coding Scheme for Content of Posts

| Category | Description | Dictionaries |
|---|---|---|
| Experience | Statements related to personal experiences (I/my as references + any verb) | LIWC:<br>First-person single pronoun,<br>Present and past verbs |
| Knowledge | Statements of opinions or ideas, factual information (non I/my as references + cognitive word) | LIWC:<br>First-person plural pronoun,<br>Second-person pronoun,<br>Impersonal pronoun,<br>Noun,<br>Article,<br>Cognitive process. |
| Inquiry | Direct questions | Question mark: "?" |
| Sociality | Supportive statements, and statements related to personal feelings and desires. | LIWC:<br>Affect.<br>Self-defined category:<br>Greeting and gratitude. |

between a new poster and a new referenced poster in other threads within 1 month before the new post was made in the ongoing thread. The strength of ties is measured along five dimensions: (1) coexistence, (2) directness, (3) recency, (4) reciprocity, and (5) frequency as described earlier. We present the measures of all the variables in the main study in Table 11.

## Data Analysis and Models

To investigate the reference relationship in an ongoing thread, we removed the first and second posts in each thread since in one thread, both the first and second posts were made to the group rather than to responders. In all, we included 431 posts in our analysis. We undertook mixed quantitative analysis. We first conducted Pearson's correlation analysis to identify highly correlated independent variables included in the study. We then used binary logistical regression analysis to investigate whether a new poster referenced a group or a responder and whether the new poster referenced a new friend or an old friend (H1, H2, and H3). Because of unequal sample sizes for variables indicating the strength of shared ties in other threads (illustrated in the following section). We conducted independent sample *t* tests to explore the association of reference relationships in the ongoing thread and the strength of ties in other threads (H4).

We constructed two binary logistic regression models. In Model 1, the dependent variable is whether a new poster referenced the group (0) or a responder (1), and the independent variables are variables indicating the new posters' engagement in the community and in the ongoing thread and the categories of content of the new posts. In Model 2, the dependent variable is whether a new poster referenced an old friend (0) or a

Table 11.Measures for Variables in the Main Study

| Variable | Measure | Code |
|---|---|---|
| **Reference relationship in the ongoing thread** | Reference to the group (0) vs. a responder (1) | |
| | Reference to an old friend (0) vs. a new friend (1) | |
| **New posters' engagement in the community** | | |
| Duration of membership in the community | Number of days since the new poster became a member | Days_Member |
| Engagement with the community | Number of threads previously initiated by the new poster in the community | Num_Threads |
| | Effectiveness of initiating threads by the new poster: ratio of Num_Threads to Days_Member | Threads_Days |
| Engagement with individuals | Number of posts previously made by the new poster in the community | Num_Posts |
| | Effectiveness of making posts by the new poster: ratio of Num_Posts to Days_Member | Posts_Days |
| | Number of likes that the new poster previously gave to specific posts in the community | Num_Likes |
| | Effectiveness of giving likes: ratio of Num_Likes to Days_Member | Likes_Days |
| **New poster's engagement in the ongoing thread** | | |
| Duration of engagement | Number of days since the new poster's first post in the ongoing thread | Days_FirstPost |
| Recency of engagement | Number of days since the new poster's most recent post in the ongoing thread | Days_RecentPost |
| Activity of engagement | Number of posts previously made by the new poster in the ongoing thread | Num_PostsOngo |
| Duration of the ongoing thread | Number of days since the ongoing thread was initiated | Num_DayThread |
| | | |

Table 11. continued

| Variable | Measure | Code |
|---|---|---|
| **Strength of ties in at least one other thread between the new poster and new referenced poster** | | |
| Coexistence | Appearing in at least one thread in 1 month before the new poster made the current post in the ongoing thread: no (0) vs. yes (1) | AppearOneThread |
| Directness | One of the two posters directly referenced the other in at least one other thread: indirect link (0) vs. direct link (1) | DirectLink |
| Reciprocity | The two posters directly referenced each other in at least one other thread: nonreciprocal (0) vs. reciprocal (1) | Reciprocity |
| Frequency | The two posters directly referenced each other in at least one other thread: one time (0) vs. multiple times (1) | Frequency |
| Recency | Number of days between when the most recent post was made by the two posters in one other thread and when the new post was made between the two posters in the ongoing thread | Recency |
| **Content of the new post in the ongoing thread** | | |
| Experience | Number of statements indicating experience | Experience |
| Knowledge | Number of statements indicating knowledge | Knowledge |
| Inquiry | Number of statements indicating inquiry | Inquiry |
| Sociality | Number of statements indicating sociality | Sociality |

new friend (1), and the independent variables are variables indicate the new posters' engagement in the community and in the ongoing thread and the categories of content of the new posts.

## Results and Discussion

Overall, the results provide substantial evidence that reference relationships in an ongoing thread are significantly associated with posters' engagement in the community and in the ongoing thread, the content of posts in the ongoing thread, and ties in other threads (Table 12). In this section, we first present the results of Pearson's correlation analysis of independent variables, and then demonstrate and discuss the results of two logistic regression models and one independent sample *t* test.

### Correlation Analysis

Multicollinearity among independent variables influences the statistical significance of each independent variable in logistic regression models that we test for this study. The independent variables indicating new posters' engagement in the community are highly correlated with each other. In particular, the three variables indicating the absolute value of new posters' engagement in the community (Num_Threads, Num_Posts, Num_Likes) are significantly and strongly correlated with each other ($p < 0.01$, $R^2 > 55\%$).

The variables indicating the length of membership, the number of posts, the number of likes given to others each has significant ($p < 0.05$) but weak correlation ($R^2 < 50\%$) with two of the three variables indicating the effectiveness of new poster's

Table 12. Summary of Hypotheses and Results

| Hypotheses | Support from Analysis | Notes |
|---|---|---|
| H1a & H2a | Partial support | New posters' engagement with the community had negative effects on the likelihood of referencing a specific responder (Model 1) |
| H1b | No support | New posters' engagement in ongoing threads had no effects on the likelihood of referencing a specific responder (Model 1) |
| H2b | Full support | New posters' engagement in ongoing threads had positive effects on the likelihood of referencing a new friend (Model 2) |
| H3a & H4a | No support | Social messages had no effects on the likelihood of referencing a specific responder (Model 1) or referencing a new friend (Model 2). |
| H3b & H4b | Partial support | Inquiry messages had no effects on the likelihood of referencing a specific responder (Model 1) or referencing a new friend (Model 2). |
| H3c | Full support | Experience messages had negative effects on the likelihood of referencing a specific responder (Model 1) |
| H4c | No support | Experience messages had no effects on the likelihood of referencing a new friend (Model 2) |
| H5 | Full support | Tie strength in other threads was positively correlated with the likelihood of referencing a new friend. |

engagement (Threads_Days, Posts_Days, and Likes_Days). The number of posts by new posters has significant ($p < 0.01$) but weak correlation ($R^2 < 40\%$) with all the three variables indicating effectiveness of new posters' engagement.

The independent variables indicating new posters' engagement in the ongoing thread are also highly correlated ($P < 0.01$). However, the independent variables indicating different types of content of new posts showed weaker correlations. The variable indicating the number of statements on knowledge is positively correlated with the variables indicating the number of statements on inquiry ($p < 0.01$, $R^2 = 15\%$). We present the results of correlation analysis in Tables 13, 14, and 15.

## Model 1

In Model 1, we investigate how new posters' engagement in the community and the ongoing thread and the content of new posts are associated with the probability of the new poster referencing the group or a responder in the ongoing thread. Our dataset includes 10 threads with 431 posts. Among the 431 posts, 170 posts (39%) were coded as "0" (referencing the group) and 261 posts (61%) were coded as "1" (referencing other responders). For all 431 posts, there are a total of 2362 statements coded.[2] Among them, 1022 statements indicating experience, 977 statements indicating knowledge, 105 statements indicating inquiry, and 258 statements indicating sociality. In total, more than 84% of statements indicate experience and knowledge. The categories of experience and knowledge are the top two categories of content of the threaded posts in the community. We present the descriptive statistics for all the variables in Model 1 in Table 16.

[2] Twenty-three statements were uncoded. We removed the uncoded statements from the study.

Table 13. Correlations of Independent Variables Indicating Engagement in the Community

| | Independent Variable | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|---|
| 1 | Days_Member | 1 | | | | | | |
| 2 | Num_Threads | .594** | 1 | | | | | |
| 3 | Num_Posts | .625** | .623** | 1 | | | | |
| 4 | Num_Likes | .606** | .657** | .921** | 1 | | | |
| 5 | Threads_Days | -.167** | .309** | 0.002 | 0.026 | 1 | | |
| 6 | Posts_Days | -0.031 | .146** | .405** | .336** | .341** | 1 | |
| 7 | Likes_Days | .095* | .256** | .478** | .517** | .188** | .815** | 1 |
| ** Correlation is significant at the 0.01 level (2-tailed). | | | | | | | | |
| * Correlation is significant at the 0.05 level (2-tailed). | | | | | | | | |

Table 14. Correlations of Independent Variables Indicating Engagement in the Ongoing Thread

| | Independent Variable | 1 | 2 | 3 | 4 |
|---|---|---|---|---|---|
| 1 | Days_FirstPost | 1 | | | |
| 2 | Days_RecentPost | .633** | 1 | | |
| 3 | Num_PostsOngo | .310** | .130** | 1 | |
| 4 | Num_DayThread | .550** | .359** | .158** | 1 |
| ** Correlation is significant at the 0.01 level (2-tailed). | | | | | |

Table 15. Correlations of Variables indicating the Content of Posts

| | Independent Variable | 1 | 2 | 3 | 4 |
|---|---|---|---|---|---|
| 1 | Experience | 1 | | | |
| 2 | Knowledge | -0.05 | 1 | | |
| 3 | Inquiry | 0.043 | .152** | 1 | |
| 4 | Sociality | 0.014 | -0.072 | 0.055 | 1 |
| ** Correlation is significant at the 0.01 level (2-tailed). | | | | | |

Table 16. Descriptive Statistics for Independent Variables in Model 1

| Independent Variable | Referenced to the group (0) vs. responder (1) | Mean | Std. Deviation | Std. Error Mean |
|---|---|---|---|---|
| Days_Member | 0 | 391.71 | 379.177 | 29.082 |
| | 1 | 300.03 | 376.836 | 23.326 |
| New poster's engagement in the community | | | | |
| Num_Threads | 0 | 40.83 | 97.198 | 7.477 |
| | 1 | 36.23 | 71.258 | 4.411 |
| Num_Posts | 0 | 1172.445 | 2201.541 | 168.851 |
| | 1 | 1877.534 | 3898.016 | 241.281 |
| Num_Likes | 0 | 608.07 | 1563.456 | 119.912 |
| | 1 | 1117.26 | 2870.179 | 177.66 |
| Threads_Days | 0 | 0.149 | 0.3 | 0.023 |
| | 1 | 0.178 | 0.276 | 0.017 |
| Posts_Days | 0 | 3.334 | 4.246 | 0.326 |
| | 1 | 6.068 | 7.595 | 0.47 |
| Likes_Days | 0 | 1.196 | 2.393 | 0.184 |
| | 1 | 3.019 | 4.587 | 0.284 |
| New poster's engagement in the ongoing thread | | | | |
| Days_FirstPost | 0 | 4.259 | 19.923 | 1.528 |
| | 1 | 13.775 | 32.737 | 2.038 |
| Days_RecentPost | 0 | 3.035 | 15.025 | 1.152 |
| | 1 | 4.364 | 14.394 | 0.896 |
| Num_PostsOngo | 0 | 0.929 | 1.819 | 0.140 |
| | 1 | 2.973 | 2.928 | 0.182 |
| Num_DayThread | 0 | 24.947 | 49.897 | 3.827 |
| | 1 | 42.387 | 51.272 | 3.174 |
| Content of the new post | | | | |
| Experience | 0 | 2.706 | 3.464 | 0.266 |
| | 1 | 2.162 | 2.886 | 0.179 |
| knowledge | 0 | 2.046 | 2.700 | 0.207 |
| | 1 | 2.218 | 2.484 | 0.154 |
| Sociality | 0 | 0.647 | 2.614 | 0.201 |
| | 1 | 0.569 | 1.111 | 0.069 |
| Inquiry | 0 | 0.271 | 0.613 | 0.047 |
| | 1 | 0.227 | 0.582 | 0.036 |

Engagement in the Community

We conducted a binary logistic regression analysis to investigate how a new poster's engagement in the community predicts whether the new poster referenced the group (0) or a responder (1). The Hosmer and Lemeshow Test for goodness of fit shows a nonsignificant chi-square, indicating that the data fit the model well ($p = 0.406$). Table 17 shows the logistic regression coefficients, Wald test, and odds ratios for each of the predictors. The coefficient of the length of membership of new posters ($B = -0.001$) is negative and significant at the $p = 0.01$ level, suggesting that as the length of membership of new posters increases, the likelihood of referencing a specific responder will decrease. The coefficient of the number of likes given per day to specific responders in the community ($B = 0.147$) is positive and significant at the $p = 0.05$ level, suggesting that as the effectiveness of giving likes to other posters increases, the likelihood of referencing a specific responder will increase.

Engagement in the Ongoing Thread

A binary logistic regression analysis with the variables indicating the new poster's engagement in the ongoing thread as independent variables fails to predict whether the new poster referenced the group or a specific responder. The Hosmer and Lemeshow Test for goodness of fit shows a significant chi-square indicating that the data did not fit the model well ($p = 0.001$).

Table 17. Logistical Regression Analysis for
Engagement in the Community

| Predictor | *B* | *S.E.* | *Wald* | *P* | *Odds* |
|---|---|---|---|---|---|
| Days_Member | -0.001 | 0.000 | 8.438 | 0.004 | 0.999 |
| Num_Posts | 0.000 | 0.000 | 1.214 | 0.271 | 1.000 |
| Num_Threads | -0.002 | 0.002 | 0.634 | 0.426 | 0.998 |
| Num_Likes | 0.000 | 0.000 | 0.000 | 0.990 | 1.000 |
| Likes_Days | 0.147 | 0.066 | 4.920 | 0.027 | 1.159 |
| Posts_Days | -0.005 | 0.036 | 0.021 | 0.886 | 0.995 |
| Threads_Days | 0.007 | 0.512 | 0.000 | 0.989 | 1.007 |
| Constant | 0.484 | 0.184 | 6.906 | 0.009 | 1.623 |

Content of Posts

A binary logistic regression analysis of the effects of the content of a new post on whether a new poster referenced the group or a responder shows significant effects. The Hosmer and Lemeshow Test for goodness of fit shows a nonsignificant chi-square, indicating that the data fit the model well ($p = 0.612$). Table 18 shows the logistic regression coefficients, Wald test, and odds ratios for each of the predictors. The coefficient of the number of statements indicating knowledge ($B = 0.092$) is positive and significant at the $p = 0.05$ level, suggesting that as the knowledge-related statements increases, the likelihood of referencing a specific responder will increase. The coefficient of the number of statements indicating experience ($B = -0.060$) is negative and significant at the $p = 0.1$ level, suggesting that as the experience-related statements increases, the likelihood of referencing a specific responder will decrease.

A Summary of Model 1

We included all the predictors that showed significant effects in the above logistical regression analysis in Model 1. The Hosmer and Lemeshow Test for goodness of fit shows a nonsignificant chi-square indicating that the data fit the model well ($p = 0.917$). Table 19 shows the logistic regression coefficients, Wald test, and odds ratios for each of the predictors. The coefficient of the length of membership of new posters ($B = -0.001$) is negative and significant at the $p = 0.01$ level, suggesting that when a poster is an older member, the likelihood of referencing a specific responder will decrease. The coefficient of the number of likes given to specific responders per day ($B = 0.175$) is positive and significant at the $p = 0.01$ level, suggesting that as a poster more actively

Table 18. Logistical Regression Analysis
for the Content of Posts

| Predictor | *B* | *S.E.* | *Wald* | *P* | *Odds* |
|---|---|---|---|---|---|
| Experience | -0.060 | 0.032 | 3.398 | 0.065 | 0.943 |
| Knowledge | 0.092 | 0.039 | 5.501 | 0.019 | 1.002 |
| Inquiry | -0.150 | 0.167 | 0.797 | 0.372 | 0.861 |
| Sociality | 0.004 | 0.054 | 0.005 | 0.942 | 1.004 |
| Constant | 0.813 | 0.165 | 24.16 | 0.000 | 2.255 |

Table 19. Logistic Regression Analysis for Model 1

| Predictor | *B* | *S.E.* | *Wald* | *P* | *Odds* |
|---|---|---|---|---|---|
| Experience | -0.065 | 0.034 | 3.732 | 0.053 | 0.937 |
| Knowledge | 0.080 | 0.042 | 3.647 | 0.056 | 1.003 |
| Days_Member | -0.001 | 0.000 | 10.645 | 0.001 | 0.999 |
| Likes_Days | 0.175 | 0.038 | 21.250 | 0.000 | 1.191 |
| Constant | 0.754 | 0.198 | 14.569 | 0.000 | 2.126 |

gives likes to others, the likelihood of referencing a specific responder will increase.

The number of statements indicating experience and knowledge had significant but weaker effects. The coefficient for the number of statements indicating experience ($B$ = -0.065) is negative and significant at the $p$ = 0.1 level, suggesting that as a post includes more experience-related statements, the likelihood of referencing a specific responder will decrease. However, the coefficient of the number of statements indicating knowledge ($B$ = 0.080) is positive and significant at the $p$ = 0.1 level, suggesting that as a post includes more knowledge-related statements, the likelihood of referencing a specific responder will increase.

In all, results for Model 1 indicate that the new poster who referenced a responder was highly engaged in interaction with individual posters by actively giving likes to those posters, and in other-focused cognitive activities demonstrated in their online discourse emphasizing knowledge rather than experience. Interestingly, the new poster who referenced a responder was more likely to be a newer member in the community. This suggests that newer members may keep an open mind and actively interact with other individuals in the community. The new poster's engagement in the ongoing thread does not have a significant influence on whether the new poster references the group or a responder.

## Model 2

In Model 2, we investigate how new posters' engagement in the community and the ongoing thread and the content of new posts are associated with the probability of the new poster referencing an old friend or a new friend in the ongoing thread. Our dataset

includes 258 posts in which a responder was referenced. Among the 258 posts, 152 posts (59%) were coded as "0" (referencing an old friend), and 106 posts (41%) were coded as "1" (referencing a new friend). For all 258 posts, there are a total of 1301 statements. Among them, 562 statements indicate experience, 532 statements indicate knowledge, 59 statements indicate inquiry, and 148 statements indicate sociality. We present the descriptive statistics for the variables in Model 2 in Table 20.

Engagement in the Community

We conducted a binary logistic regression analysis to investigate how a new posters' engagement in the community predicts whether the new poster referenced an old friend (0) or a new friend (1). The Hosmer and Lemeshow Test for goodness of fit shows a nonsignificant chi-square, indicating that the data fit the model well ($p = 0.272$). Table 21 shows the logistic regression coefficients, Wald test, and odds ratios for each of the predictors. The coefficient for the length of membership of new posters ($B = 0.002$) is positive and significant at the $p = 0.01$ level, suggesting that when a poster is an older member, the likelihood of referencing a new friend will increase. The coefficient for the number of threads initiated by the new poster in the community per day ($B = -2.612$) is negative and significant at the $p = 0.05$ level, suggesting that as a poster more actively initiates threads, the likelihood of referencing a new friend will decrease. The coefficient for the number of likes given to specific responders is positive and significant at the $p = 0.1$ level, suggesting that as a poster gives more likes to others, the likelihood of referencing a new friend will increase.

Table 20. Descriptive Statistics for Independent Variables in Model 2

| Independent Variable | referenced old friend (0) vs. new friend (1) | Mean | Std. Deviation | Std. Error Mean |
|---|---|---|---|---|
| Engagement in the community | | | | |
| Days_Member | 0 | 205.213 | 328.529 | 26.388 |
| | 1 | 432.912 | 389.225 | 38.539 |
| Num_Threads | 0 | 25.845 | 59.611 | 4.788 |
| | 1 | 49.392 | 77.020 | 7.626 |
| Num_Posts | 0 | 1252.123 | 3310.183 | 265.880 |
| | 1 | 2852.407 | 4545.128 | 450.035 |
| Num_Likes | 0 | 747.426 | 2440.585 | 196.033 |
| | 1 | 1662.157 | 3364.820 | 333.167 |
| Threads_Days | 0 | 0.213 | 0.322 | 0.026 |
| | 1 | 0.114 | 0.122 | 0.012 |
| Posts_Days | 0 | 5.965 | 7.053 | 0.567 |
| | 1 | 6.103 | 6.419 | 0.636 |
| Likes_Days | 0 | 2.892 | 4.105 | 0.330 |
| | 1 | 3.113 | 4.357 | 0.431 |
| Engagement in the ongoing thread | | | | |
| Days_FirstPost | 0 | 9.000 | 21.057 | 1.714 |
| | 1 | 20.708 | 43.663 | 4.241 |
| Days_RecentPost | 0 | 1.689 | 8.450 | 0.688 |
| | 1 | 8.217 | 19.484 | 1.892 |
| Num_PostsOngo | 0 | 3.874 | 2.883 | 0.235 |
| | 1 | 1.717 | 2.502 | 0.243 |
| Num_DayThread | 0 | 38.019 | 48.122 | 3.878 |
| | 1 | 49.132 | 55.174 | 5.359 |
| Content of the posts | | | | |
| Experience | 0 | 2.390 | 3.130 | 0.252 |
| | 1 | 1.830 | 2.467 | 0.240 |
| Knowledge | 0 | 1.805 | 2.304 | 0.186 |
| | 1 | 2.396 | 2.696 | 0.262 |
| Sociality | 0 | 0.708 | 1.283 | 0.103 |
| | 1 | 0.368 | 0.760 | 0.074 |
| Inquiry | 0 | 0.201 | 0.541 | 0.044 |
| | 1 | 0.264 | 0.637 | 0.062 |

Table 21. Logistical Regression Analysis for
Engagement in the Community

| Predictor | *B* | *S.E.* | *Wald* | *P* | *Odds* |
|---|---|---|---|---|---|
| Days_Member | 0.002 | 0.001 | 10.371 | 0.001 | 1.002 |
| Num_Threads | -0.006 | 0.005 | 1.519 | 0.218 | 0.994 |
| Num_Posts | 0.000 | 0.000 | 0.022 | 0.882 | 1.000 |
| Num_Likes | 0.001 | 0.000 | 0.211 | 0.084 | 1.001 |
| Threads_Days | -2.612 | 1.164 | 5.031 | 0.025 | 0.073 |
| Posts_Days | 0.063 | 0.056 | 1.247 | 0.264 | 1.065 |
| Likes_Days | 0.051 | 0.090 | 0.326 | 0.568 | 1.053 |
| Constant | -0.865 | 0.255 | 11.489 | 0.001 | 0.421 |

Engagement in the Ongoing Thread

Next, our logistic regression analysis includes the variables indicating the new poster's engagement in the ongoing thread as the independent variables. The Hosmer and Lemeshow Test for goodness of fit shows a nonsignificant chi-square, indicating that the data fit the model well ($p = 0.443$). Table 22 shows the logistic regression coefficients, Wald test, and odds ratios for each of the predictors. Two predictors showed positive and significant effects at the $p = 0.05$ level. As the number of days since the new poster's most recent post in the ongoing thread was made ($B = 0.239$) increases, the likelihood of referencing a new friend will increase. As the number of days since the ongoing thread was initiated ($B = 0.005$) increases, the likelihood of referencing a new friend will increase.

Content of Posts

A binary logistic regression analysis of the effects of the content of a new post on whether a new poster referenced an old friend or a new friend shows significant effects. The Hosmer and Lemeshow Test for goodness of fit shows a nonsignificant chi-square indicating that the data fit the model well ($p = 0.662$). Table 23 shows the logistic regression coefficients, Wald test, and odds ratios for each of the predictors. At the $p = 0.05$ level, two predictors showed significant effects. The coefficient for the number of statements indicating knowledge ($B = 0.116$) is positive, suggesting that as the number of the knowledge-related statements increases, the likelihood of referencing a specific responder will increase. However, the coefficient for the number of statements indicating inquiry is negative, suggesting that as the number of the inquiry-related statements

Table 22. Logistic Regression Analysis for Engagement in the Ongoing Thread

| Predictor | *B* | *S.E.* | *Wald* | *P* | *Odds* |
|---|---|---|---|---|---|
| Days_FirstPost | -0.001 | 0.007 | 0.046 | 0.829 | 0.999 |
| Days_RecentPost | 0.239 | 0.127 | 3.548 | 0.031 | 1.273 |
| Num_PostsOngo | -0.090 | 0.061 | 2.129 | 0.145 | 0.914 |
| Num_DayThread | 0.005 | 0.003 | 2.815 | 0.043 | 1.005 |
| Constant | -0.442 | 0.215 | 4.248 | 0.039 | 0.643 |

Table 23. Logistic Regression Analysis for Content of New Post

| Predictor | *B* | *S.E.* | *Wald* | *P* | *Odd* |
|---|---|---|---|---|---|
| Experience | -0.068 | 0.049 | 1.949 | 0.163 | 0.934 |
| Knowledge | 0.116 | 0.054 | 4.585 | 0.032 | 1.122 |
| Inquiry | -0.379 | 0.150 | 6.352 | 0.012 | 0.684 |
| Sociality | 0.199 | 0.218 | 0.833 | 0.361 | 1.220 |
| Constant | -0.323 | 0.216 | 2.226 | 0.136 | 0.724 |

increases, the likelihood of referencing a responder will decrease.

A Summary of Model 2

We included all the predictors that showed significant effects in the binary logistical regression analyses in one logistic regression model. The Hosmer and Lemeshow Test for goodness of fit shows a nonsignificant chi-square, indicating that the data fit the model well ($p = 0.488$). Table 24 shows the logistic regression coefficients, Wald test, and odds ratios for each of the predictors. The coefficient for the length of membership of new posters ($B = 0.002$) is positive and significant at the $p = 0.01$ level, suggesting that when a poster is an older member, the likelihood of referencing a new friend will increase. The coefficient for the number of likes given to specific responders ($B = 0.001$) is positive and significant at the $p = 0.05$ level, suggesting that as a poster gives more likes to others, the likelihood of referencing a new friend will increase. Three predictors showed positive and significant effects at the $p = 0.1$ level. As the number of days since the new poster made the most recent post in the thread ($B = 0.223$) increases, the likelihood of referencing a new friend will increase. As the number of days since the thread was initiated ($B = 0.004$) increases, the likelihood of referencing a new friend will increase. As the number of statements indicating knowledge ($B = 0.078$) increases, the likelihood of referencing a new friend will increase.

To sum up, a new poster's membership status, engagement with individuals in the community, engagement in the ongoing thread, and the content of posts indicating knowledge significantly predict whether the new poster referenced an old friend or a new friend. The new poster was more likely to reference a new friend when the poster was an

Table 24. Logistic Regression Model for Model 2

| Predictor | *B* | *S.E.* | *Wald* | *P* | *Odds* |
|---|---|---|---|---|---|
| Days_Member | 0.002 | 0.001 | 8.597 | 0.003 | 1.002 |
| Threads_Days | -1.031 | 0.76 | 1.839 | 0.175 | 0.357 |
| Num_Likes | 0.001 | 0.004 | 2.085 | 0.049 | 1.005 |
| Days_RecentPost | 0.223 | 0.131 | 2.89 | 0.089 | 1.249 |
| Num_DayThread | 0.004 | 0.003 | 1.709 | 0.091 | 1.004 |
| Knowledge | 0.078 | 0.059 | 1.791 | 0.081 | 1.082 |
| Inquiry | 0.180 | 0.250 | 0.522 | 0.470 | 1.198 |
| Constant | -1.328 | 0.322 | 17.01 | 0.000 | 0.265 |

older member in the community, gave more likes to other posters, and had been involved in the ongoing thread longer. When referencing a new friend, the new poster demonstrated more other-focused cognitive activities.

## Ties in Other Threads

In this analysis, we investigate how the strength of ties in other threads between the new poster and new referenced posters are associated with the probability of the new poster referencing an old friend or a new friend in the ongoing thread. Our dataset includes 258 posts. Among the 258 posts, 152 posts (59%) were coded as "0" (referencing an old friend), and 106 posts were coded as "1" (referencing a new friend). We present the frequencies for variables in Table 25. Among new posters, 67.50% of new posters referencing an old friend and 79.20% of new posters referencing a new friend appeared in at least 1 other thread in one recent month before the poster posted in the ongoing thread. Among new posters who appeared in one other thread with the new referenced posters, more than 80% of new posters referencing an old friend and referencing a new friend had direct, reciprocal, frequent, and recent (within 4 days) reference relationships with the new referenced posters in the other thread. This indicates that new posters and new posters in ongoing threads had strong reference relationships in other threads in the community regardless of whether the new referenced poster was a new friend or an old friend. We present the descriptive statistics for variables in Table 26. An independent sample *t* test shows that the two groups were significantly different with regard to the variable indicating whether the new poster and new referenced poster appeared in at least one thread in the most recent month (AppearOneThread; see Table

Table 25. Frequency of Independent Variables indicating Tie Strength

| Variable | | | Referenced new poster: old friend (0) vs. new friend (1) | |
|---|---|---|---|---|
| | | | 0 | 1 |
| AppearOneThread | 0 | Count | 49 | 22 |
| | | % | 32.50% | 20.80% |
| | 1 | Count | 102 | 84 |
| | | % | 67.50% | 79.20% |
| | Total | Count | 151 | 106 |
| Directlink | 0 | Count | 1 | 2 |
| | | % | 1.20% | 3.20% |
| | 1 | Count | 84 | 60 |
| | | % | 98.80% | 96.80% |
| | Total | Count | 85 | 62 |
| Reciprocity | 0 | Count | 1 | 2 |
| | | % | 1.20% | 3.20% |
| | 1 | Count | 84 | 60 |
| | | % | 98.80% | 96.80% |
| | | Count | 85 | 62 |
| Frequency | 0 | Count | 14 | 10 |
| | | % | 16.70% | 17.20% |
| | 1 | Count | 70 | 48 |
| | | % | 83.30% | 82.80% |
| | | Count | 84 | 58 |

Table 25. continued

| Variable | | | Referenced new poster: old friend (0) vs. new friend (1) | |
|---|---|---|---|---|
| | | | 0 | 1 |
| Recency | 0 | Count | 21 | 23 |
| | | % | 21.00% | 30.30% |
| | 1 | Count | 37 | 18 |
| | | % | 37.00% | 23.70% |
| | 2 | Count | 16 | 14 |
| | | % | 16.00% | 18.40% |
| | 3 | Count | 3 | 2 |
| | | % | 3.00% | 2.60% |
| | 4 | Count | 4 | 5 |
| | | % | 4.00% | 6.60% |
| | 5 | Count | 3 | 1 |
| | | % | 3.00% | 1.30% |
| | 6 | Count | 3 | 1 |
| | | % | 3.00% | 1.30% |
| | 7 | Count | 1 | 2 |
| | | % | 1.00% | 2.60% |
| | 10 | Count | 0 | 2 |
| | | % | 0.00% | 2.60% |
| | 11 | Count | 1 | 0 |
| | | % | 1.00% | 0.00% |
| | 12 | Count | 0 | 1 |
| | | % | 0.00% | 1.30% |
| | 13 | Count | 2 | 0 |
| | | % | 2.00% | 0.00% |
| | 14 | Count | 1 | 1 |
| | | % | 1.00% | 1.30% |
| | 15 | Count | 2 | 1 |
| | | % | 2.00% | 1.30% |
| | 18 | Count | 1 | 0 |
| | | % | 1.00% | 0.00% |
| | 20 | Count | 1 | 2 |
| | | % | 1.00% | 2.60% |
| | 21 | Count | 1 | 1 |
| | | % | 1.00% | 1.30% |

Table 26. Descriptive Statistics for Variables Indicating Tie Strength

| Variable | Referenced an old friend (0) vs. a new friend(1) | N | Mean | Std. Deviation | Std. Error Mean |
|---|---|---|---|---|---|
| AppearOneThread | 0 | 151 | 0.671 | 0.471 | 0.038 |
| | 1 | 106 | 0.792 | 0.408 | 0.040 |
| DirectLink | 0 | 103 | 0.709 | 0.390 | 0.038 |
| | 1 | 86 | 0.816 | 0.457 | 0.049 |
| Recency | 0 | 100 | 3.54 | 6.016 | 0.602 |
| | 1 | 76 | 3.592 | 6.134 | 0.704 |
| Reciprocity | 0 | 85 | 0.988 | 0.109 | 0.012 |
| | 1 | 62 | 0.968 | 0.178 | 0.023 |
| Frequency | 0 | 85 | 0.833 | 0.375 | 0.041 |
| | 1 | 62 | 0.828 | 0.381 | 0.050 |

27). For this categorical variable, appearing in one specific thread is coded as "1," and not appearing in that thread is coded as "0." Results show that the new poster who referenced a new friend in the ongoing thread was more likely to appear in at least one other thread with the new referenced poster (the two posters previously made posts in at least one other thread in the community; $p = 0.033$) and have direct links with the new referenced poster ($p = 0.086$). These results indicate that for a new poster, a new friend is a "new" friend in the ongoing thread, but an "old" friend in other threads. The two posters participate in at least one other thread in the community and have direct reference relationships in the shared thread.

Table 27. Independent Sample *T* Test for Tie Strength

| Variable | *t* test for Equality of Means | | | | | | |
|---|---|---|---|---|---|---|---|
| | *t* | *df* | *Sig.* | Mean Difference | Std. Error Difference | 95% Confidence Interval of the Difference | |
| | | | | | | Lower | Upper |
| AppearOneThread | 2.150 | 256.000 | 0.033 | -0.121 | 0.057 | 0.010 | 0.233 |
| DirectLink | 1.725 | 187.000 | 0.086 | -0.106 | 0.062 | -0.228 | 0.015 |
| Reciprocity | -0.864 | 145.000 | 0.389 | 0.021 | 0.024 | -0.067 | 0.026 |
| Frequency | -0.089 | 140.000 | 0.929 | 0.006 | 0.064 | -0.133 | 0.122 |
| Recency | 0.056 | 174.000 | 0.955 | -0.052 | 0.923 | -1.770 | 1.874 |

# GENERAL DISCUSSION

Social learning is a shared phenomenon. Therefore, properties of social relationships play a crucial role in social learning. Previous research on consumer online communities has focused primarily on who is whose friend. Online communities allow users to create a list of friends. Previous research has primarily focused on how users are influenced by their online friends in terms of attitudes and behaviors (Ellison, Steinfield, & Lampe, 2007; Habibi, Laroche, & Richard, 2014; Wilcox & Stephen, 2013). A simple friend count is an insufficient metric for user influence (Green, 2008; Trusov et al., 2010). Trusov et al. (2010) proposed that the influentials were not those who have the most friends, but those whose log-in frequency positively predicts their friends' log-in frequency.

To more fully understand how consumers learn through online social networks, we argue that it is important to understand specific features of a relationship, including the strength of that relationship. Our study undertakes a relational view to explore the reference relationships between posters within threads, which are key driving factors for an enduring social learning process in online communities. Specifically, in our preliminary study, we find that one poster is more likely to be referenced by others in an ongoing thread when the poster's engagement in the community was high (Proposition 1) and the content of the new post by the poster in the ongoing thread focused on advanced learning (Proposition 2 ) and sociality (Proposition 3). Likewise, in our main study, we

discover that one poster is more likely to reference a responder, especially a new friend in the ongoing thread, when the new poster's engagement in the community (H1) and the ongoing thread (H2) is high and the content of the poster's new post in the ongoing thread focused on advanced learning (H3). We also find that new posters and new referenced posters had prior ties in other threads. New posters who referenced new friends in the ongoing thread had stronger ties with the new referenced posters than did new posters who referenced old friends (H4).

Our study enhances the present understanding of the mechanism of consumer learning in online communities by examining the pattern of referencing relationships in online learning networks, or Who-Referenced-Whom networks. From an ecological point of view, posters in one community are connected by engaging in a variety of activities in the community such as giving likes to others, initiating threads, and referencing others in specific threads. A new poster's reference behavior in a specific thread is influenced by this new poster's network established in the community. We identify which engagement activities, personal ties, and categories of content of posts motivate new posters' decisions about referencing. Specifically, we explore new posters' behavior in referencing the group or a specific responder and in referencing a new friend or an old friend.

Our study demonstrates that the strength of ties between posters plays different roles at the macro and micro levels in reference relationships in ongoing threads. At the macro level, posters participate across threads in a given community and through that participation form ties with the community and with individual posters. Posters' engagement in the community is demonstrated through engaging in a variety of activities

such as initiating threads, making posts, and giving likes to other posters. These activities help posters form ties with the community and with specific posters. At the micro level, posters learn in one thread demonstrated by the levels of the content of posts (e.g., experience and knowledge) and establish reference relationships with a group or specific responders. Ties established in the community may drive posters to make new posts to particular responders, especially those with whom new posters have had ties in other threads but with whom they have not had ties in the ongoing thread. This illustrates how ties are built through the transitivity of relationships across threads. Granovetter (1973) argued that in social networks, small scale interaction (interaction with individuals) can translate into large scale patterns (a network of relationships), which in turn influence small groups. Likewise, in an online community, interactions between two posters in one thread can translate into a reference network in the community. From this perspective, reference relationships drive the learning process in ongoing threads, but also help build ties in the community. In ongoing threads, reference relationships established in other threads motivate new posters to make posts to responders with the content of posts focusing on knowledge and inquiry. Such relationships may strengthen ties between new posters and new referenced posters, which will positively influence their reference relationships in the next threads.

Reference relationships are influenced not only by posters' ties established in the community demonstrated by their historical activities (e.g., engagement across threads in the community), but also revealed by the cognitive activities of participating in ongoing threads demonstrated by the content of new posts. New posters who referenced responders showed cognitive processes with more focus on others or the group (using

second-or third-person and first-person plural pronouns in their posts) rather than on themselves (using first-person pronouns). In a learning environment, other-or group-focused content (e.g., "you") and cognitive content are more likely to have information that can be directly used by others or the group than are self-focused (e.g., "I") content. These cognitive activities reflect an advanced step in the social learning process: knowledge generation. From this perspective, reference relationships advance learning in ongoing threads.

Specifically, referencing a new friend in an ongoing thread reflects a new poster's intention to initiate a new conversation with this new friend. Such an activity may encourage the cultivation of fresh ideas and self-discovery in the learning process. Our study shows that the new poster who references a new friend is more likely to have shared ties with the new friend in other threads. This suggests that the strength of shared ties plays a significant role in facilitating the dissemination of new information in ongoing threads. In our study we assess strength of shared ties in threads other than the ongoing threads (intensity of social interactions between posters) along five dimensions: existence, directness, frequency, reciprocity, and recency. Our study provides empirical evidence that new posters who referenced responders (either new friends or old friends) had strong shared ties with new referenced posters in at least one other thread in the month before the two posters interacted in the ongoing threads. Specifically, the mean of existence and directness of shared ties had higher value for new posters referencing new friends. This suggests that the strength of shared ties influences a new poster's intention to initiate a new conversation.

The sociality-related content of posts has been ignored by previous research in

consumer social learning. Our preliminary study shows that sociality-related content is positively associated with both the establishment of reference relationships and tie strength in ongoing threads. This suggests that sociality-related messages provide posters an opportunity for collaborating and extend the opportunities for learning. However, in our main study, sociality-related content did not show significant effects on reference relationships in ongoing threads. This suggests that social content is a significant factor influencing whether a post is referenced, but not a significant factor in whether a new post references a specific responder rather than the group. That is, sociality-related content may be a signal that the poster welcomes interaction, but that signal does not prioritize interaction with self or the group.

Among all the statements coded in our study, more than 80% of statements indicate experience and knowledge. This shows that in health-related online communities, posters focus on the exchange of knowledge and experience. As we discussed earlier, experience sharing and knowledge generation are two important steps in online learning. Posters share experience as they identify problems and generate knowledge as they reflect, refine, and explore problems (Jayanti & Singh, 2009). Knowledge generation can be a driver or a consequence of posters' high engagement in the learning process demonstrated by reference relationships in the ongoing thread.

# IMPLICATIONS

Consumers can benefit significantly from online social learning (e.g., empowering decision making; Jayanti & Singh, 2009). It is important for marketers to facilitate consumers' social learning through online communities. Our study suggests the factors related to reference relationships in an ongoing thread, including (1) posters' engagement in community activities and (2) posters' engagement in the ongoing thread, (3) ties between the new poster and new referenced poster in other threads, and (4) the content of posts reveal progressive cognitive activity advancing social learning.

General online communities such as Facebook provide users opportunities for a variety of social activities (e.g., event creation and invitation, photo sharing). However, topic-specialized online communities, especially health-related online communities, only provide a limited number of functions for social activities. Our study shows that sociality is an important factor for consumer learning. Consumers learn as they build and develop ties with one another. We suggest that topic-specialized online communities should increase and diversify social tools to facilitate consumers' learning process.

Mapping the content of communication in virtual communities alone does not provide a clear understanding of the social learning process and thus a clear path to intervention. For example, in some threads, posters progress through the early stages of social learning, such as identifying problems (problem-focused experience) and refining

problems (questions), but do not advance to the higher stages of social learning such as acquiring knowledge (comment) and transferring knowledge to action (action-focused experience). Without knowing the ties between posters, it is difficult to suggest an intervention that could extend the social learning process. However, if we view social learning as the result of the establishment and development of ties between members, we can assess which parts of the social structure of the thread may be inhibiting the continuation of the social learning process. For example, if the posters' engagement in the community is the problem, tools to demonstrate posters' engagement (e.g., rating for credibility) should be considered in the design of online communities. If the content of posts is the problem (e.g., lacking sociality-related statements), forum moderators may play a role in facilitating discussion within threads to extend the social learning process.

Crowdsourcing is a process in which businesses acquire ideas and content about their products or services from a large group of people especially in online communities instead of from their employees and suppliers (Crowdsourcing, n.d.). Our study shows that new posts in which responders are referenced contained more knowledge-related content. Such posts can be used as a valuable source for crowdsourcing. We suggest that online communities may develop a set of analytical tools for the data related to quoting. For example, in one thread, key measures may include the number of quotations, the number of unique posts that are quoted, and the number of unique posters whose posts are quoted.

Indentifying influentials is a key strategy for viral marketing. Suppliers of online social networking services such as Facebook and Google are developing algorithms to identify and target advertisements to influentials. Social media marketers have not taken

full advantage of the social components of social media. Many brands have focused primarily on individuals rather than relationships among individuals and the learning consequences of those relationships. Our study suggests that social media marketers should identify the most influential "teachers" in online communities, who are referenced most frequently in other posts, and the most hard-working "students," who reference other posters most frequently. We argue that the most influential teachers are most likely to influence others' opinions or actions in the community, whereas the most hard-working students are most likely to need specific information about products or services. In online dialogues, both "teachers" and "students" are engaged in advanced learning (e.g., inquiry and knowledge generation), providing opportunities for dissemination of knowledge about new products or advanced functions of existing products.

Social media marketers should collect and analyze data about the online behaviors of "teachers" and "students" in order to get insights into the learning process. Key questions that should be answered may include what discussion topics are most likely to involve the most teachers and students? How do teachers and students interact across multiple threads in the community? And who is most likely to initiate a new conversation (teachers or students). For online community design, we suggest that marketers add a "Follower" function. Users could then follow those who are teachers, and their influence on followers and friends could be measured.

# FUTURE RESEARCH

We propose and test a framework for dynamism of online social learning by building on the previous literature on social learning, community, and social network (especially tie strength). Our study focuses on the predictors of two types of reference relationships: referencing a group or a responder, and referencing a new friend or an old friend. This research suggests some directions for future research.

Our study has shown the strong effects of poster's engagement, the content of posts, and strength of shared ties on reference relationships in health-related online communities. Research is needed to investigate the factors that influence reference relationships in other types of online communities such as brand communities and general communities (e.g., Twitter, Facebook). Especially, research is needed to explore the factors that influence the probability of retweeting. Retweeting is an important function in Twitter, indicating the behavior that users re-post someone else's tweets. Re-posting others' tweets is an action similar to quoting others' posts and could extend the social learning process.

Our study has shown the usefulness of sets of variables for the measurement of strength of shared ties in threads other than ongoing threads. More research is needed to identify common and unique variables for the measurement of strength of shared ties in different types of online communities.

Our study measures behavioral engagement in an online community using variables such as the number of posts previously made in the community. Research using surveys to investigate the psychological factors (e.g., trust, credibility) underlying online behavioral engagement could provide additional insights about motivations for the observed patterns of behavior.

REFERENCES


Adamic, L. A., Zhang, J., Bakshy, E., & Ackerman, M. S. (2008, April). Knowledge sharing and Yahoo Answers: Everyone knows something. *Proceedings from the 17th International Conference on World Wide Web* (pp. 665–674). New York, NY: ACM.

Ahern, T. C., Peck, K., & Laycock, M. (1992). The effects of teacher discourse in computer-mediated discussion. *Journal of Educational Computing Research*, *8*(3), 291–309.

Argote, L. (2012). *Organizational learning: Creating, retaining and transferring knowledge*. New York, NY: Springer.

Arguello, J., Butler, B. S., Joyce, E., Kraut, R., Ling, K. S., Rosé, C., & Wang, X. (2006). Talk to me: Foundations for successful individual-group interactions in online communities. *Proceedings from the SIGCHI Conference on Human Factors in Computing Systems* (pp. 959–968). New York, NY: ACM.

Baer, M. (2012). Putting creativity to work: The implementation of creative ideas in organizations. *Academy of Management Journal*, *55*(5), 1102–1119.

Baller, R. D., & Richardson, K. K. (2009). The "dark side" of the strength of weak ties: The diffusion of suicidal thoughts. *Journal of Health and Social Behavior*, *50*(3), 261–276.

Bateman, P. J., Gray, P. H., & Butler, B. S. (2011). Research note — The impact of community commitment on participation in online communities. *Information Systems Research*, *22*(4), 841–854.

Bandura, A. (1971). *Social learning theory*. New York, NY: General Learning Press.

Borgatti, S. P., & Cross, R. (2003). A relational view of information seeking and learning in social networks. *Management Science. 49*(4), 432–445.

Breiger, R., Carley, K., & Pattison, P. (Eds.). (2003). *Dynamic social network modeling and analysis: Workshop summary and papers*. Washington, DC: National Academies Press.

Brooks, R., & Meltzoff, A. N. (2005). The development of gaze following and its relation to language. *Developmental Science, 8*(6), 535–543.

Brown, A. L., & Palincsar, A. S. (1989). Guided, cooperative learning and individual knowledge acquisition. In *Knowing, learning, and instruction: Essays in honor of Robert Glaser* (pp. 393–451). Oxford, UK: Routledge.

Brown, J. J., & Reingen, P. H. (1987). Social ties and word-of-mouth referral behavior. *Journal of Consumer Research*, 350–362.

Bullen, M. (2007). Participation and critical thinking in online university distance education. *International Journal of E-Learning & Distance Education, 13*(2), 1–32.

Burke, M., Marlow, C., & Lento, T. (2009, April). Feed me: Motivating newcomer contribution in social network sites. *Proceedings from The SIGCHI Conference on Human Factors in Computing Systems* (pp. 945–954). New York, NY: ACM.

Burt, R. S. (2004). Structural holes and good ideas. *American Journal of Sociology*, *110*(2), 349–399.

Calder, B. J., Malthouse, E. D., & Schaedel, U. (2009). An experimental study of the relationship between online engagement and advertising effectiveness. *Journal of Interactive Marketing*, *23*(4), 321–331.

Celen, B., Kariv, S., & Schotter, A. (2010). An experimental test of advice and social learning. *Management Science*, *56*(10), 1687–1701

Chen, F. C., & Wang, T. C. (2009). Social conversation and effective discussion in online group learning. *Educational Technology Research and Development, 57*(5), 587–612.

Cross, R., & Sproull, L. (2004). More than an answer: Information relationships for actionable knowledge. *Organization Science*, *15*(4), 446–462.

Crowdsourcing. (n. d.) In Merriam-Webster online. Retrieved from http://www.merriam-webster.com/dictionary/crowdsourcing

Donelle, L., & Hoffman-Goetz, L. (2008). An exploratory study of Canadian Aboriginal online health care forums. *Health Communication*, *23*(3), 270–281.

Duan, W., Gu, B., & Whinston, A. (2008). The dynamics of online word-of-mouth and product sales: An empirical investigation of the movie industry. *Journal of Retailing, 84*(2), 233–242.

Eichhorn, K. C. (2008). Soliciting and providing social support over the Internet: An

investigation of online eating disorder support groups. *Journal of Computer Mediated Communication*, *14*(1), 67–78.

Ellison, N. B. (2007). Social network sites: Definition, history, and scholarship. *Journal of Computer Mediated Communication*, *13*(1), 210–230.

Ellison, N. B., Steinfield, C., & Lampe, C. (2007). The benefits of Facebook "friends:" Social capital and college students' use of online social network sites. *Journal of Computer Mediated Communication*, *12*(4), 1143–1168.

Faraj, S., & Johnson, S. L. (2011). Network exchange patterns in online communities. *Organization Science*, *22*(6), 1464–1480.

Fehr, E., & G ächter, S. (2000). Fairness and retaliation: The economics of reciprocity. *Journal of Economic Perspectives*, *14*(3), 159–181.

Feld, S. L. (1981). The focused organization of social ties. *American Journal of Sociology*, *86*(5), 1015 – 1035.

Friedkin, N. (1980). A test of structural features of Granovetter's strength of weak ties theory. *Social Networks*, *2*(4), 411–422.

Gilbert, E., Karahalios, K., & Sandvig, C. (2008). The network in the garden: an empirical analysis of social media in rural life. *Proceedings from the SIGCHI Conference on Human Factors in Computing Systems* (pp. 1603–1612). New York, NY: ACM.

Glasgow, R. E., Christiansen, S. M., Kurz, D., King, D. K., Woolley, T., Faber, A. J., ... & Dickman, J. (2011). Engagement in a diabetes self-management website: Usage patterns and generalizability of program use. *Journal of Medical Internet research*, *13*(1). doi: 10.2196/jmir.1391.

Granovetter, M. S. (1973). The strength of weak ties. *American Journal of Sociology*, *78*(6), 1360–1380.

Granovetter, M. (1983). The strength of weak ties: A network theory revisited. *Sociological Theory*, *1*(1), 201–233.

Green, H. (2008, October 6). Google: Harnessing the power of cliques. *Business Week*, 50. Retrieved from http://www.thefreelibrary.com/GOOGLE: HARNESSING THE POWER OF CLIQUES-a01611662657.

Greene, J. A., Choudhry, N. K., Kilabuk, E., & Shrank, W. H. (2010). Online social networking by patients with diabetes: A qualitative evaluation of communication with Facebook. *Journal of General Internal Medicine*, *26*(3), 287–92.

Gutter, M. S., Garrison, S., & Copur, Z. (2010). Social learning opportunities and the financial behaviors of college students. *Family and Consumer Sciences Research Journal, 38*, 387–404.

Habibi, M. R., Laroche, M., & Richard, M. O. (2014). Brand communities based in social media: How unique are they? Evidence from two exemplary brand communities. *International Journal of Information Management*, *34*(2), 123–132.

Hansen, M. T., Mors, M. L., & Løvås, B. (2005). Knowledge sharing in organizations: Multiple networks, multiple phases. *Academy of Management Journal*, *48*(5), 776–793.

Hazan, C., & Shaver, P. (1987). Romantic love conceptualized as an attachment process. *Journal of Personality and Social Psychology*, *52*(3), 511–524.

Henri, F. (1992). Computer conferencing and content analysis. *Collaborative Learning through Computer Conferencing*, *90*, 117–136.

Hsi, S., & Hoadley, C. M. (1997). Productive discussion in science: Gender equity through electronic discourse. *Journal of Science Education and Technology*, *6*(1), 23–36.

Huang, K., & Li, S.Y. (2007). The influence of eWOM on virtual consumer communities: Social capital, consumer learning, and behavioral outcomes. *Journal of Advertising Research*, *47*(4), 485–495.

Ingram, A. L., & Hathorn, L. G. (2004). Methods for analyzing collaboration in online. In T. S. Roberts (Ed.), *Online collaborative learning: Theory and practice* (pp. 215–241). Hershey, PA: Information Science Publishing.

Jayanti, R.K., & Singh, J (2010). Pragmatic learning theory: An inquiry action framework for distributed consumer learning in online communities. *Journal of Consumer Research, 36*(6), 1058–1081.

Joyce, E., & Kraut, R. E. (2006). Predicting continued participation in newsgroups. *Journal of Computer-Mediated Communication*, *11*(3), 723–747.

Kozinets, R. V. (1997). "I Want to Believe": A nethnography of the 'X-Philes' subculture of consumption. *Advances in Consumer Research, 24*, 470–475.

Kozinets, R. V., & Handelman, J. (1998). Ensouling consumption: A netnographic exploration of the meaning of boycotting behavior. *Advances in consumer research*, *25*(1), 475–480.

Kozinets, R.V. (1998) On netnography: Initial reflections on consumer research investigations of cyberculture. *Advances in Consumer Research, 25*, 366–371.

Kozinets, R. V. (1999). E-tribalized marketing?: The strategic implications of virtual communities of consumption. *European Management Journal*, *17*(3), 252–264.

Kraatz, M. S. (1998). Learning by association? Interorganizational networks and adaptation to environmental change. *Academy of Management Journal*, *41*(6), 621–643.

Krackhardt, D. (1992). The strength of strong ties: The importance of philos in organizations. *Networks and Organizations: Structure, Form, and Action*, *216*, 239.

Kreijns, K., & Kirschner, P. A. (2001). The social affordances of computer-supported collaborative learning environments. *Proceedings from Frontiers in Education Conference, 2001. 31st Annual* (Vol. 1, pp. T1F-12). Hoboken, NJ: IEEE.

Kreijns, K., Kirschner, P. A., & Jochems, W. (2002). The sociability of computer-supported collaborative learning environments. *Educational Technology & Society*, *5*(1), 8–22.

Lakhani, K. R., & Von Hippel, E. (2003). How open source software works: "Free" user-to-user assistance. *Research Policy*, *32*(6), 923–943.

Lave, J., & Wenger, E. (1990). *Situated learning: Legitimate peripheral participation*, Cambridge, UK: Cambridge University Press.

Levin, D. Z., & Cross, R. (2004). The strength of weak ties you can trust: The mediating role of trust in effective knowledge transfer. *Management Science*, *50*(11), 1477–1490.

Lin, N., Dayton, P. W., & Greenwald, P. (1978). Analyzing the instrumental use of relations in the context of social structure. *Sociological Methods & Research*, *7*(2), 149–166.

Louch, H. (2000). Personal network integration: Transitivity and homophily in strong-tie relations. *Social Networks*, *22*(1), 45–64.

Messner, M., Linke, M., & Eford, A. (2011). Shoveling tweets: An analysis of the microblogging engagement of traditional news organizations. *Proceedings from 12th International Symposium for Online Journalism.* Austin, Texas: University of Texas at Austin

Morales, M., Mundy, P., Delgado, C. E., Yale, M., Messinger, D., Neal, R., & Schwartz, H. K. (2000). Responding to joint attention across the 6-through 24-month age period and early language acquisition. *Journal of Applied Developmental Psychology*, *21*(3), 283–298.

Moretti, E. (2011). Social learning and peer effects in consumption: Evidence from movie sales. *Review of Economic Studies*, *78*(1), 356–393

Newman, M., Barabasi, A. L., & Watts, D. J. (2011). *The structure and dynamics of networks*. Princeton, NJ: Princeton University Press.

Newman, M. L., Pennebaker, J. W., Berry, D. S., & Richards, J. M. (2003). Lying words: Predicting deception from linguistic styles. *Personality and Social Psychology Bulletin*, *29*(5), 665–675.

Newman, M. L., Groom, C. J., Handelman, L. D., & Pennebaker, J. W. (2008). Gender differences in language use: An analysis of 14,000 text samples. *Discourse Processes*, *45*(3), 211–236.

Park, D. H., Lee, J., & Han, I. (2007). The effect of on-line consumer reviews on consumer purchasing intention: The moderating role of involvement. *International Journal of Electronic Commerce*, *11*(4), 125–148.

Pawan, F., Paulus, T. M., Yalcin, S., & Chang, C. (2003). Online learning: Patterns of engagement and interaction among in-service teachers. *Language Learning & Technology*, *7*(3), 119–140.

Pena-Shaff, J. B., & Nicholls, C. (2004). Analyzing student interactions and meaning construction in computer bulletin board discussions. *Computers & Education*, *42*(3), 243–265.

Pennebaker, J. W., Mayne, T. J., & Francis, M. E. (1997). Linguistic predictors of adaptive bereavement. *Journal of Personality and Social Psychology*, *72*(4), 863.

Pennebaker, J. W., & Graybeal, A. (2001). Patterns of natural language use: Disclosure, personality, and social integration. *Current Directions in Psychological Science*, *10*(3), 90-93.

Pennebaker, J. W., Mehl, M. R., & Niederhoffer, K. G. (2003). Psychological aspects of natural language use: Our words, our selves. *Annual review of psychology*, *54*(1), 547–577.

Phelps, C., Heidl, R., & Wadhwa, A. (2012). Knowledge, networks, and knowledge networks: A review and research agenda. *Journal of Management*, *38*(4), 1115–1166.

Pool, I. (1980). Comment on Mark Granovetter's 'The strength of weak ties: A network theory revisited'. *Proceedings from the International Communications Association in Acapulco*. Washington, DC: ICA.

Rafaeli, S., & Ariel, Y. (2008). Online motivational factors: Incentives for participation

and contribution in Wikipedia. *Psychological Aspects of Cyberspace: Theory, Research. 24*, 1913–1928.

Ransbotham, S., & Kane, G. C. (2011). Membership turnover and collaboration success in online communities: Explaining rises and falls from grace in Wikipedia. *MIS Quarterly-Management Information Systems*, *35*(3), 613.

Redcay, E., Kleiner, M., & Saxe, R. (2012). Look at this: The neural correlates of initiating and responding to bids for joint attention. *Frontiers in Human Neuroscience*, *6*, 169.

Riegner, C. (2007) Word of mouth on the web: The impact of Web 2.0 on consumer purchase decisions. *Journal of Advertising Research*, *47*(4), 436–447.

Schaefer, C., Coyne, J. C., & Lazarus, R. S. (1981). The health-related functions of social support. *Journal of Behavioral Medicine*, *4*(4), 381–406.

Schau, H. J., Muñiz Jr, A. M., & Arnould, E. J. (2009). How brand community practices create value. *Journal of Marketing*, *73*(5), 30–51.

Scott, J. (2012). *Social network analysis*. Thousand Oaks, CA: SAGE Publications Limited.

Shi, X., Adamic, L. A., & Strauss, M. J. (2007). Networks of strong ties. *Physica A: Statistical Mechanics and its Applications*, *378*(1), 33–47.

Slatcher, R. B., Vazire, S., & Pennebaker, J. W. (2008). Am "I" more important than "we"? Couples' word use in instant messages. *Personal Relationships*, *15*(4), 407–424.

Small, S., & Supple, A. (2001). Communities as systems: Is a community more than the sum of its parts. In A. Booth & A. C. Crouter (Eds.), *Does it take a village? Community effects on children, adolescents, and families* (pp. 161–174). Mahwah, NJ: Lawrence Erlbaum Associates.

Sze, P. (2008). Nurturing teacher leadership qualities through computer conferencing: A study of teachers' online interaction patterns. In A. Li & L. Shiu (Eds.), *Developing teachers and developing schools in changing contexts* (pp. 419). Hong Kong, CN: Chinese University Press.

Tahmincioglu, E. (2008). Facebook Friends as Job References. Retrieved from http://www.msnbc.msn.com/id/26223330

Tausczik, Y. R., & Pennebaker, J. W. (2010). The psychological meaning of words: LIWC and computerized text analysis methods. *Journal of language and Social Psychology*, *29*(1), 24–54.

Trusov, M., Bodapati, A. V., & Bucklin, R. E. (2010). Determining influential users in internet social networks. *Journal of Marketing Research*, *47*(4), 643–658.

Tumasjan, A., Sprenger, T. O., Sandner, P. G., & Welpe, I. M. (2010). Predicting elections with Twitter: What 140 characters reveal about political sentiment. *Proceedings from the International AAAI Conference on Weblogs and Social Media*, *10*, 178–185.

Uzzi, B., & Lancaster, R. (2003). Relational embeddedness and learning: The case of bank loan managers and their clients. *Management Science*, *49*(4), 383–399.

Walther, J. B. (1992). Interpersonal effects in computer-mediated interaction: A relational perspective. *Communication Research*, *19*(1), 52–90.

Walther, J. B. (1996). Computer-mediated communication impersonal, interpersonal, and hyperpersonal interaction. *Communication Research*, *23*(1), 3–43.

Wang, P., Robins, G., Pattison, P., & Lazega, E. (2013). Exponential random graph models for multilevel networks. *Social Networks*, *35*(1), 96–115.

Wasserman, S., & Galaskiewicz, J. (Eds.). (1994). *Advances in social network analysis: Research in the social and behavioral sciences*. Oaklands, CA: Sage Publications.

Wegerif, R. (1998). The social dimension of asynchronous learning networks. *Journal of Asynchronous Learning Networks*, *2*(1), 34–49.

Wellman, B. (2008). Review: The development of social network analysis: A study in the sociology of science. *Contemporary Sociology*, *37*, 221–222.

Wenger, E. (1999) Communities of practice: The key to a knowledge strategy. *Knowledge Directions*, *1*(2), 48–63.

Wenger, E. (2006). Communities of Practice: A Brief Introduction. Retrieved from http://www.ewenger.com/theory/index.htm

White, D. R., & Houseman, M. (2002). The navigability of strong ties: Small worlds, tie strength, and network topology. *Complexity*, *8*(1), 72–81.

Wilcox, K., & Stephen, A. T. (2013). Are close friends the enemy? Online social networks, self-esteem, and self-control. *Journal of Consumer Research*, *40*(1), 90–103.

Wise, K., Hamman, B., & Thorson, K. (2006). Moderation, response rate, and message interactivity: Features of online communities and their effects on intent to participate. *Journal of Computer Mediated Communication*, *12*(1), 24–41.